

What will happen now to India?

Will the bomb that killed Mr Rajiv Gandhi last week be the trigger that breaks the subcontinent into smaller pieces? Both India and the rest of us must hope not.

THE assassination last week of Mr Rajiv Gandhi is not just a tragedy, but a danger for us all. Among poor countries, India has the richest promise. Not for nothing is it known as the world's most populous democracy. That the electoral process (during which Gandhi was killed) has survived for more than four decades in such diversity of interest and allegiance — ethnic, religious, economic and geographical, for example — is a marvel. It owes as much to the emphasis laid by the leaders of the independence movement of the 1930s and 1940s on the interdependence of urban and rural India as to the vigour with which Mr Rajiv Gandhi's grandfather, Jawaharlal Nehru, embraced the doctrine of secular democratic government. Can it survive a third assassination of a Gandhi? And what, otherwise, will be the consequences?

The observations that India is a contrast of wealth and poverty, and of modernity and backwardness, are familiar but not trite. The contrasts are often shocking to people from elsewhere, puzzled that peasant villages still limp along without elementary sanitary technology in a country that builds its own nuclear power stations and orbital space rockets. The justification of the persistence of these contrasts is a kind of trickle-down theory of social and economic development — crumbs spill from rich tables, people who cut their teeth on building nuclear power stations eventually turn their attention to village latrines. There has been some proof, in the past few decades, that the theory works, but not quickly enough to assuage regional passions and discontents. The greatest irony is in the Punjab, where the green revolution first brought a degree of prosperity — and then the separatist movement that became the indirect cause of Rajiv Gandhi's mother's assassination.

Salutary shock

The gloomy view is that the latest assassination is not so much an occasion that may prove to be a trigger for the break-up of India, but instead a symptom of a process already begun. In that case, there is a chance that the shock the occasion will provoke may have a salutary effect. Gandhi, after all, is but one of more than a hundred people killed so far in the course of this general election. Tamils, Sikhs and even the Hindu fundamentalists of poverty-stricken Bihar may now be persuaded that the goals behind the communal violence of the past few years are not prizes for which lives (often other people's) are worth sacrificing, but are empty prizes. All-union India may be virtually unmanageable, but a splintered India would be even nearer chaos.

Even within India, the benefits of union are too often overlooked. Not the least of these is what union has done for India's intellectual life. Since independence, the centre has created throughout provincial India universities in the image of India's best, both by the efforts of enlightened public servants at the University Grants Committee and by the private foundation of institutions as different as the Tata Institute at Bombay and the Raman Institute at Bangalore, both (among others in India) now research centres of international repute. In short, the centre has made possible the creation in India of a respected intellectual cadre whose chief misfortune is that so many of its members choose the brighter opportunities overseas. Even the Gupta business at Chandigarh, with which the university has not yet dealt effectively, may be taken as a measure of the respect accorded to intellectual pursuits; would people take such trouble to simulate true scholarship if it were otherwise?

Secularism

Socially, the benefits of union are even greater. There are, for example, more Muslims (75 million plus) in India than in any of the predominantly Muslim states except Indonesia, yet for most of India's independence, tensions familiar elsewhere have been abated by the secularism of the state. Would that civility persist if India were broken into pieces? The centre has also given even the poor a sense that civic rights are there to be exercised, even in the face of a bureaucracy whose bumbledom is by now legendary. The union has been less successful at spreading the benefits of technical change in rural India. Indirectly, that may have been another cause of last week's tragedy.

From the outside, India's union has been an even greater benefit. Since independence, there has been a single voice to speak for getting on for 1,000 million people. Although this collective message has often irritated people in the West, on questions such as the Nuclear Non-Proliferation Treaty for example, even those who have been most maddened by India's prevarications and casuistry internationally will readily acknowledge how much less comfortable it would be if there were a dozen new voices on the international scene preoccupied with regional issues — Kashmir, perhaps, or the Tamil population of Sri Lanka. That is yet another reason why India must hang together. In a society still tolerant of the notion that there are leaders and the led, that is a challenge for the intellectual community of India. The goal must be to make the poor of India prosperous. □



\$100 billion loan

There would be sense in regarding seriously the Soviet Union's need of funds.

THE First International Sakharov Conference last week in Moscow invaluabley flushed out more data about the consequences of the Chernobyl accident as well as exploring somewhat grander themes — the management of the global energy economy, for example. But the occasion also showed how much has changed since Sakharov died eighteen months ago. Then, the proceedings of such a conference, still important, would have been of absorbing and immediate interest. But, last weekend, the more important meeting was not in Moscow, but at Cambridge, Massachusetts, where a small group of economists from the Soviet Union was hammering out, with American advice, a plan to revitalize the Soviet economy. The understanding seems to be that, if a workable plan can be agreed, the United States (and the other industrialized members of the Group of Seven) will provide the Soviet Union with credit. And if not, they will not.

In the annals of diplomacy, this meeting will rank as one of the strangest. The economists sitting at Harvard University were not government officials, let alone diplomats. Neither they nor their interlocutors can tell whether an economic plan that seems viable in Massachusetts will be acceptable in Moscow or, if it were, whether it could be carried through effectively. And there is no way of telling what weight the opinions of those looking over the economists' shoulders will carry with the US and other governments. The potential for an accommodation rests on the mutual appreciation by the two sides of each others' difficulties. The West has grown to know Mr Mikhail Gorbachev, and does not want him to fall, last year's weakness over the economy and the Baltic states notwithstanding. And Gorbachev needs credits — a sum of \$100,000 million has been mentioned — or he will have a brief future.

There are two snags. One is that Gorbachev might win the promises he needs and yet be unable to push through the reforms that would be their precondition; then the Soviet Union would fall over the edge on which it has been teetering for months. The other, unspoken but more immediate, is that the West cannot afford the investment the Soviet Union needs; by any standards, \$100,000 million is a lot of money. Western governments, already strapped for capital (see *Nature* 347, 315; 1990) and in recession, are most of all concerned to know whether, and when, they can promise that their own people's prosperity will be restored. Few of them will be inclined to think that there are votes in giving precedence to Soviet needs.

But in the long run, the calculation may be wrong. Even in the short run, \$100,000 million would not be as onerous as it seems. Because largesse on such a scale directed towards the Soviet Union would be spent mostly in the West, the extra business might be just the fillip required to lift the West out of recession. There are flies in that ointment, to be sure, as the condition of reunited Germany now shows; inflationary

pressures have grown, interest rates are high and people in what used to be West Germany feel impoverished. (So they do in the East, but for different reasons.) Helping the Soviet Union on a significant scale would certainly leave an economic mark. So, too, would significant help for the really poor countries of the world. The worry is that the yearning for a return of the old pattern of prosperity will mean that opportunities are passed up. Certainly, that is what Gorbachev is saying. □

Does industry corrupt?

How often does a company's attempt to pervert a scientific presentation succeed?

THE received wisdom that there is hardly any way in which a commercial company can tempt a true researcher from the narrow path of rectitude is not, of course, an absolute truth. Everybody can think of a few exceptions to the rule of virtue which, by the flagrancy of their subordination of data to the exigencies of the market, are so exceptional that they seem to be a proof that the rule is generally kept. But out-and-out breaches of the principle that science is science and industry is what commerce can make of it are probably less important than more subtle pressures to bend the truth.

Dr Peter J. Quesenberry, the distinguished haematologist from the University of Virginia at Charlottesville, has a cautionary tale to tell in the current issue of *Experimental Hematology*, of which he is editor-in-chief. For some years, the annual meeting of the American Society of Hematologists has apparently been preceded by an event called "super-Friday" at which commercial companies invite the principal speakers at the scientific meeting to give broad accounts of what they will be saying at the technical sessions that will follow. That the practice is a public service is easily understood, especially in a field so in the thick of rapid revolution as haematology because of the importance in medicine of newly discovered lymphokines, growth factors and even chromosome mapping. The companies do the decent thing, and pay those who speak on super-Friday.

Quesenberry's experience at last December's Boston meeting is nevertheless a little chilling. He explains that he was asked by his host (a pharmaceutical company) to rehearse his intended talk for super-Friday over an "elegant dinner", when it was suggested that he should omit from his presentation some slides showing that certain growth factors, although blessed with therapeutic potential, appeared also to stimulate the growth of tumour cells. On the face of things, those concerned appeared to be worried that open talk of potential side-effects might dim the tale of promise there would otherwise be to tell. Quesenberry's own reaction to the request can be inferred from the way in which he has retold the story. Two questions remain. What is to become of drug companies so stupid that they are unable to recognize the value of early warning of side-effects? And — more worrying — is everybody put in Quesenberry's position as honourable as he? □

Chernobyl effects not as bad as feared

■ Study finds no cancer or birth defects

■ Two key populations omitted, though

Vienna

THE first international study on the medical effects of the Chernobyl nuclear accident has found that, contrary to media reports, the 1986 accident did not result in any measurable radiological effects on the health of the local population. The only health effects of the accident were related to psychological effects such as stress, the study concluded.

The year-long study, presented last week at a four-day scientific meeting of the International Atomic Energy Agency (IAEA) in Vienna, was carried out at the request of the Soviet Union between March 1990 and January 1991. Its purpose was to assess the effects of the accident on the 825,000 inhabitants of the three most contaminated republics, the Ukraine, Byelorussia and the Russian Federation.

The researchers who carried out the study admit that it had several limitations. Because they were limited in both time and money, the visiting researchers relied on spot checks and statistical analysis of Soviet results instead of attempting to verify Soviet measurements from scratch.

Furthermore, the researchers did not look at two populations thought to be most at risk from radiation: the emergency clean-up crews known as 'liquidators' and the more than 100,000 residents who were evacuated in 1986 from the prohibited 30-km zone around the reactor.

Fred Mettler of the University of New Mexico, a member of the study team, said that the liquidators would have been a more valuable, if logistically difficult, group to study, but that the research team was restricted to observing the population that the Soviet government had requested. Soviet experts at the meeting said that they too were very interested in studying the effects of radiation on the liquidators, but that it would be difficult for a number of reasons, including insufficient data on the health of the liquidators before the accident.

The team members did say, however, that they could state with confidence that there is no known increase observed in any radiation-related illnesses, such as leukaemia, birth defects or thyroid cancer, for the populations of villages near the reactor that had been contaminated by radiation. Only in the case of thyroid cancer did researchers consider it likely that a significant, if small, number of excess cases might occur in the next five to ten years.

The study is seen as the most convincing evidence yet that the Soviets had not been covering up any disastrous health consequences of the accident — charges that had

been widely reported in the media (see *Nature* 351, 4; 1991). The study was sponsored not only by IAEA, but also by six other United Nations agencies, and the scientists on the team stressed that they carried out their measurements independently of IAEA.

Not everyone at the meeting accepted the team's conclusions, however. Representatives from the Ukraine objected to what they called the "optimism" of the study. The situation, at least in the Ukraine, is "definitely worse" than described in the study, said Viktor Bar'yakhtar, vice-chairman of the Ukrainian Academy of Sciences in Kiev. He said that there are statistically significant increases observed in both pregnancy abnormalities, such as stillbirths and vaginal bleeding, and thyroid disorders in children.

Environmental organizations also criticized the study. Greenpeace issued a statement saying that the study was "fully insufficient" and has grave weaknesses.

On the other hand, participants familiar with the data on the long-term radiological effects of the atomic bombs dropped on Hiroshima and Nagasaki said it is not at all surprising that no excess cancers have been reported yet at Chernobyl because it is just too early — and the Japanese experience does not point to a large increase in cancer even decades after the initial exposure.

Itsuzo Shigematsu, director of the Radiation Effects Research Foundation in Hiroshima, said a five-year study on survivors of the atomic blasts would not have identified many excess cancers either, despite the much higher radiation doses received. Even after 45 years, said Shigematsu, just 700 excess cancers had been identified among about 110,000 survivors.

Indeed, the study implied that the Soviet government's overreaction to the accident poses a greater danger to the population than the radiation. The study found that the purposely conservative Soviet measurements placed the radiation dose absorbed by the population at two to three times its actual value. Politicians in Moscow and the republics then used these data to justify resettling hundreds of thousands of people.

According to study team member Terence Lee, a professor of psychology at University of Surrey who assessed the psychological effects of the accident, moving people out of the region now is medically unwarranted and would be a "disaster" for their health, as the trauma of relocation would be "colossal".

But the Soviet Union seems intent on continuing the resettlements. According to

a new law passed on 20 May, Moscow promises to provide relocation costs and new housing to anyone living in areas where the population is affected by a certain minimum radiation dose. Lee argues that the dose, which the law sets as an average annual radiation dose of 0.5 rem above natural background, is an "absurdly low threshold". Furthermore, Lee said, the move sets a dangerous precedent for other countries with high natural radiation levels. If Britain followed the Soviet guidelines, for example, it would "have to evacuate half of Cornwall".

Viktor Gubanov, deputy chairman of the Soviet Union Council of Ministers' State Committee on the Elimination of the Consequence of Chernobyl, says that the government will relocate at least 218,000 people between 1990 and 1992.

Incorrect information and a lack of scientific understanding of radiation have combined with political opportunism to create a huge psychological problem in the Chernobyl area, the study team said. The study found the psychological effects "wholly disproportionate" to the biological significance of the radiation.

The results "shout at you so loud," Lee said, that a study was not really needed to see how bad they are. Every little illness in the affected areas is being attributed to radiation, says Lee. "If a child gets a nosebleed, it is considered a hospital case," he says.

The study showed that, despite the absence of immediate danger, 72 per cent of adults living in the so-called "danger zone" wanted to move out.

Several members of the study team said their work was complicated by a reluctance among Soviet scientists to share all their data. "All this talk about scientific collaboration is nonsense," Lee said. The study, he said, was carried out in an atmosphere of "colossal tension" between Western and Soviet researchers — a factor that study participants fear will affect the accuracy and usefulness of future health studies in the Soviet Union.

The main problem at the outset, Lee said, was simply getting a look at the Soviets' data. Researchers would talk at meetings about their results, but they refused to give copies of their papers or summaries of their data to study team members. Mettler said it took delicate negotiations to persuade the Soviets to cooperate. "Many of them said, 'I've worked many years to get this data, and now you're going to run off and publish it'." Mettler had to promise not to copy their data, but just to take notes on it for use in the study.

Because of the public distrust of Soviet science, the scientists at the meeting saw it as essential that the researchers who performed the survey be invited to return to the Soviet Union in order to deliver its results. "If we rely now on official Soviet media to announce the results," said Friedrich Steinhäusler of the University of Salzburg, "they will just be met by more distrust."

Steven Dickman

Japan courts US on joint project

Tokyo & Washington

OFFICIALS from Japan's Ministry of International Trade and Industry (MITI) were in Washington last week to sound out the possibility of US participation in a major government-industry project MITI plans to launch next year: a programme to develop a new generation of computers that will mimic the human brain.

The New Information Processing Technology (NIPT) project, known informally as the sixth-generation computer project, will replace Japan's fifth-generation computer project after it ends next March.

When MITI announced that previous project in the early 1980s, it caught the attention of the rest of the developed world and spawned a series of competing efforts in the United States and Europe, such as the Alvey project in the United Kingdom.

That fifth-generation effort, however, which consumed about ¥50,000 million (\$370 million) over the past ten years, has fallen far short of its aims and has not led to a Japanese takeover of the world's computer industry, as some Western governments feared. So this time around, MITI is hoping to attract foreign — and especially US — expertise. Not surprisingly, US government officials have been wary about allowing US assistance to MITI's new project.

Officials of MITI's Industrial Electronics Division last week visited the White House Office of Science and Technology Policy, the

Department of Commerce and the State Department to try to encourage US participation. US officials are unwilling to reveal specifics of the meetings, but the basic question that must be answered, one official said, is: "What's in it for us?" The advantages to Japan of having collaboration from IBM or AT&T Bell Labs are clear; the advantages to the United States are not so obvious.

Tension over the project first arose last year. MITI made a number of discreet overtures to US researchers and representatives of US companies about whether they would be interested in collaboration on the project. For instance, while visiting Japan, Bell Labs president Ian Ross was asked about cooperation on the project during a conversation on a completely different subject.

But such approaches have annoyed the US government, which insists that, because this is a MITI-sponsored project, any contact must be made through official channels. In December, the Japanese agreed not to deal directly with US researchers or companies and to negotiate collaboration on the project under the guidelines of the 1988 US-Japan Science and Technology Agreement.

Those negotiations could take a long time. The only other major project to come under the agreement's umbrella is MITI's proposed Intelligent Manufacturing System (IMS) project (see *Nature* 347, 320; 1990). MITI has been seeking US support for IMS for more than a year, and so far the two countries

have done no more than agree on the "terms of reference" for a feasibility study that has not yet begun.

NIPT will focus on massively parallel computer processing and, in particular, neural network computers and optical computers. And MITI is expected to spend about \$30-\$40 million a year over ten years, says Shunichi Amari of Tokyo University, a member of a committee of more than 60 industrialists and academics that recently completed a two-year study of brain-like computers for the ministry.

Taizo Nishikawa of MITI's Industrial Electronics Division says that unlike the fifth-generation project, which concentrated on one objective, the sixth-generation project will be a research programme consisting of several projects by a number of different research groups both in Japan and overseas. Previous MITI projects have been like attempts to "climb Mount Fuji", he says, but the sixth-generation project will be "like climbing Yatsugatake", a range of eight peaks in Japan.

The next step in the US-Japan negotiations on the project will come in July, when the two countries meet in Tokyo for working-level talks on the science and technology agreement. No decision on US participation is expected before MITI submits a budget request for the first year of the project in late August.

David Swinbanks & Robert Pool

The private sector offers an alternative

WHILE the world's attention focuses on MITI's next generation computer project (see above), a group of Japanese academics and industrialists have been trying to launch an alternative to the sixth generation computer project with private sector funding.

Their initiative, called the International Institute of Novel Computing (IINC), may never get off the ground. But the very fact that some top Japanese companies have been prepared to put money into the project shows a growing movement in Japanese industry to look for alternatives to MITI projects, which, although often seen in the West as a source of Japan's economic might, are viewed in a very different light by Japanese industry.

The IINC initiative was launched last year by the Japan Technology Transfer Association — a non-profit organization funded largely by private industry — with the backing of nine private companies, including Hitachi, Nissan and Nippon Steel Corporation.

The project is led by Hideo Aiso of Keio University and includes many academics who are also involved in MITI's sixth generation project.

The nine companies each donated ¥1.8 million (\$13,000) to fund a feasi-

bility study of various types of next-generation computing on which IINC might focus, including parallel, optical and neural computing — key targets of MITI's project. When the initiative was announced last September, officials of the Technology Transfer Association said they were confident that many more companies would join the scheme (see *Nature* 347, 217; 1990).

But the expected contributors have not materialized, and academics participating in the project suspect that MITI may be to blame. One MITI official, for instance, said "we are very annoyed by IINC". He explained, "They are confusing people because many of the people participating in IINC are also participating in NIPT. I am not saying they should not be allowed to do what they want to do, but it is annoying and confusing." Simply letting companies know that MITI is "annoyed" is probably enough to kill donations without any further action.

Why has such an alternative project arisen? Although no researchers in Japanese private industry or MITI's national laboratories are prepared to say so publicly, many are dissatisfied with the ministry's big national projects.

The projects are inflexible and set rigid

goals that must be achieved no matter what the changing circumstances, researchers say. An example is the fifth generation computer project, which set out in 1982 to build in ten years a specific type of parallel computer with 1,000 microprocessors running on logic programming. By the mid-1980s, it was clear that alternative approaches might be better, but changes could not be made. IINC promoters hope to establish a longer-term project with more flexible goals.

So why does industry continue to participate in MITI's projects? The reasons are complex and include the regulatory powers MITI has over industry. One key factor is that during the early post-war era when Japanese industry was in a shambles, MITI did much to revitalize and provide support for industry. Senior executives in Japan's companies still remember those days, and in Japanese society, where obligations for past favours always remain strong, it is very hard for companies to reject MITI proposals. However, as the senior management of Japanese companies undergoes a generational change in the near future, many predict that MITI's power will wane.

D.S.

Polluters pay at auction

Washington

THE US Environmental Protection Agency (EPA) announced plans last week for setting up an official market in pollution rights for sulphur dioxide emissions. The increased efficiency in reducing emissions that would result from such a market system, says one economist who has studied such systems, could save US companies as much as \$3,000 million a year as they work to meet the provisions of the 1990 Clean Air Act.

Under the act, by the year 2000, US industry must cut total sulphur dioxide emissions to 10 million tons less than the 1980 level. Sulphur dioxide, produced mostly by coal- and oil-fired electrical generation plants, is a major cause of acid rain.

Because, for the purposes of clearing up acid rain, it does not matter which sulphur dioxide producers cut emissions, the EPA plans to let a free market govern that choice. Each current producer will receive a certain number of sulphur dioxide "allowances", where one allowance gives the right to emit one ton of sulphur dioxide a year. The number of allowances given to a source will depend on a variety of factors, including their existing emission rates.

Sulphur dioxide emitters can then trade their allowances. If, for instance, one plant can cut its emissions by one ton for \$1,000 while it would cost a second plant \$2,000, the second plant might offer \$1,500 for an allowance from the first plant.

The first would cut its emissions by a ton, sell an allowance to the second, and make \$500 on the deal; the second would save \$500 over the cost of cutting emissions itself. The effect of such a scheme is to make sure that the emission reduction is achieved at the lowest possible cost, notes Robert Hahn, an economist at the American Enterprise Institute in Washington.

In addition to the trading of allowances between companies, the EPA will offer some allowances at auction and a smaller number for direct sale at the fixed price of \$1,500 per allowance. The direct sales will be open to anyone, utility or private speculator, on a first-come, first-served basis, except that the EPA will allow independent power producers that are just starting up to get the first chance to buy the allowances.

These independent producers are smaller electricity-generating companies that normally sell power wholesale to utilities; the direct sale provisions are designed to give them the emission rights they need to begin operation.

When EPA phased out lead in gasoline, a market in lead allowances saved refiners about \$200 million a year, Hahn says, and the savings for sulphur dioxide emissions could be 10 to 15 times as great.

Robert Pool

NIH lose a legal shield

Washington

THE fate of handful of research primates removed from a Maryland laboratory — a case that began a decade ago as a minor irritation for the National Institutes of Health (NIH) — has snowballed into a Supreme Court loss for the agency that may have reverberations for decades. The high court ruled last week that NIH had wrongly derailed a lawsuit by animal rights group People for the Ethical Treatment of Animals (PETA), by having the case moved to a friendly federal court where it was dismissed.

In the decision, the court ruled that NIH had unfairly taken advantage of a pre-Civil-War law designed to protect federal tax collectors from irate local officials. The law stipulates that any suit against "any officer of the United States . . . or person acting under him" can be removed from potentially partisan local courts to the comparatively reliable federal courts. In the instance of the PETA suit, which sought to force NIH to turn over the last of the "Silver Spring monkeys" to activists, NIH lawyers, citing the law, had the suit moved from a New Orleans district court to the federal court for Louisiana. The federal court, citing "sovereign immunity" statutes that make it difficult to sue the government, threw the case out.

In its unanimous decision, the Supreme Court ruled that the law was meant specifically to protect individuals, not agencies. NIH had argued that as the director of NIH is a federal officer, the entire agency is a "person working under him" [actually "her" at this moment], and should not be sued in local

court. The Supreme Court described this as "a rather tortured reading of the language" and dismissed it.

PETA, which claimed a "tremendous victory" in the case, will now return to the New Orleans court to try again. This time it may sue Tulane University, which now holds the animals, rather than NIH. If the court decides to hear the case, it will be the first time that a judge will consider the merits of PETA's argument that the animals are being held in violation of cruelty laws. "Historically, we've been locked out of State and Federal courts," says PETA cofounder Alex Pacheco. "Now we have a crack in that wall."

NIH officials fear that the implications may go even further than the Silver Spring monkeys. The ruling appears to allow any suit against NIH to be heard in a local court, where it may not be automatically dismissed for "lack of standing", the argument usually used in federal court to halt such cases.

Although the government can always appeal against decisions to higher courts, it may now be forced to expend its resources defending itself in lower courts where it can no longer simply request a removal.

With the nationwide precedent of a Supreme Court ruling, such protracted hearings "could happen in any state", says Louis Sibal, head of NIH's animal research office. "This could engender a lot of frivolous lawsuits." One NIH lawyer, speaking on the condition of anonymity, warns that "this case is much more important for other issues than it is for the Silver Spring monkeys."

Christopher Anderson

UNIVERSITY FACULTIES

NRC proposes end to forced retirement

Washington

MOST faculty members would retire of their own accord before they reach the age of 70, even if the current mandatory retirement laws were repealed, according to a new study by the National Research Council (NRC). Professors who decide to stay past the age of 70 are generally still productive and should not be forced to leave, NRC said last week in calling for an end to special legislation that allows universities to retire tenured professors over 70.

Congress passed legislation in 1986 that outlawed age discrimination and mandatory retirement, but included an exemption for university professors. University officials had argued that the combination of tenure and no mandatory retirement would essentially guarantee professors lifetime employment, regardless of their competence. Congress asked NRC to examine the issue, to advise it on whether to extend the exemption when it automatically expires in 1994. NRC recommended that Congress drop the exemption, and for-

bid mandatory retirement on campuses. It produced data from more than 3,200 institutions indicating that the majority of tenured faculty now choose to retire before 70 and would continue to do so if current rules were lifted.

In place of retirement regulations, NRC recommended special retirement incentive programmes beginning at the age of 50 to increase the number of faculty positions open to young professors.

The cover of the NRC report was itself the object of some comment. It shows five branches of ivy on a brick wall, with four of them obviously dead and the fifth carrying a few meagre leaves in fall colours. After seeing the cover, one of the NRC committee members asked "if it was a Rorschach test," recalls Harriet Morgan, one of the editors of the report. No, Morgan says, "We just looked at a lot of generic university photos and this one seemed very colourful."

Christopher Anderson

Ending Mandatory Retirement for Tenured Faculty, National Academy Press, 1991.

India to reprocess fuels

New Delhi

INDIA'S Atomic Energy Commission has said it is ready to reprocess spent fuel from reactors of other countries as a commercial service to earn foreign exchange.

India has two fuel reprocessing plants, one in Bombay and another at Tarapur, about 100 km north of Bombay. A third plant is nearing completion near Madras. According to P. K. Iyengar, the AEC chairman, the country has sufficient reprocessing capacity to undertake work for parties abroad.

The commission would accept spent fuel from research reactors in other countries on condition that those countries would take back the wastes. Iyengar said that scientists at about 125 research reactors, mostly in developing countries, would like to have the plutonium separated from the spent fuel. He has proposed to the International Atomic Energy Agency that India could provide this service to its member countries. According to Iyengar, because the reprocessing work is manpower intensive, it would cost much less in India than in the West.

Commercial reprocessing is just one aspect of India's aggressive strategy to enter the international nuclear market, which is monopolized by Europe and North America. Over the years, the commission has built up indigenous capability for the total nuclear cycle from uranium mining to reprocessing. Iyengar said India would also like to export research reactors, related technologies such as isotope production and consultant services that are not objectionable from the point of nuclear proliferation.

Negotiations are at present going on with Egypt, Syria and some other countries. Iyengar expects that, in the future, the International Atomic Energy Agency would be able to purchase Indian-made nuclear equipment and other products for use in member countries under its technical assistance programme.

The Department of Atomic Energy is not the only agency that has set its eyes on export. The Indian Space Research Organization is planning to set up an autonomous corporation to take care of its business in India and abroad. The profits will be ploughed back into the organization.

About half of the space hardware needed by the organization — including rocket engines and propellers — is now supplied by some 100 Indian companies that have been closely associated with India's space programme from the beginning. The proposed corporation expects to help these companies to capture a share of the global space market.

As a first step towards making its presence felt internationally, the organization has made a bid for the supply of a communications satellite to the International Maritime Satellite Organisation. The space research organization is already selling rocket propellants to Canada and some components to

Europe's Ariane. Marketing of satellite images is one aim of the proposed corporation. Pictures obtained from the Indian Remote Sensing Satellite are already in demand in neighbouring countries and a second satellite is going up next month.

According to space commission chairman U. R. Rao, India could do well in the interna-

INFRARED ASTRONOMY

Keeping cool in space

London

ASTRONOMERS at the Royal Observatory, Edinburgh (ROE), believe that a simple engineering trick could greatly extend the life of the next generation of orbiting infrared telescopes. By replacing the liquid helium used to cool infrared space telescopes with a passive system that simply radiates heat out into space, telescopes could last for ten years, rather than a matter of months, says John Davies of ROE.

Astronomers from Europe, the United States and Japan met in Edinburgh last week to discuss the new idea, and to consider plans to build a large, passively cooled infrared space telescope for launch early next century.

Infrared astronomy is a difficult and expensive business. Because the atmosphere absorbs infrared light, telescopes must either be built on the top of mountains, or be launched into space.

An infrared space telescope must be kept at a very low temperature so that infrared

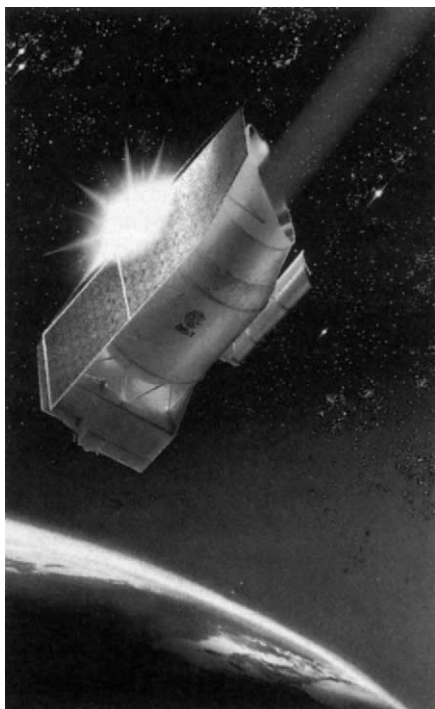
emissions from the telescope itself do not trigger the telescope's detectors. In the past, this has been achieved by insulating the telescope and cooling it cryogenically down to about 10 K using liquid helium. But the helium evaporates quickly — the pioneering Infrared Astronomy Satellite (IRAS), for example, launched in 1983, ran out of helium after only ten months in orbit.

Tim Hawarden, of ROE, turned this approach on its head in a design proposal he submitted to the European Space Agency (ESA) last year. He removed the insulation, and surrounded the telescope with a series of nested radiators (described as "cans within cans" by Davies), designed to carry heat away from the base of the telescope towards its aperture, and out into space. Hawarden's idea was not taken up by ESA, but ROE astronomers have since refined the design, which they described at last week's meeting.

Davies says that this system could cool an infrared space telescope down to 40 K. This is too warm to observe at the longer infrared wavelengths, but he says that such a telescope would work well over a wavelength band of 1–30 micrometres. "There are plenty of bits of interesting astronomy you can do in this window," Davies says. And without the need to carry a cryogenic cooling system, telescopes could be larger, giving better resolution. It should be possible to launch a telescope with a mirror 2.5 metres in diameter, says Davies. (IRAS was built with a 60-cm mirror.) The detectors on board infrared space telescopes must be cooled to near absolute zero, but Davies says this could be achieved by using a small mechanical cooling system as the detectors themselves are only about the size of a thumbnail.

Davies would like to see an international collaboration, involving Europe, the United States and Japan, to build a large infrared space telescope. But this will have to wait for several years. ESA's decision not to accept Hawarden's original proposal is not due to any obvious technical problems with the design, but rather because the agency wishes to examine the results from its Infrared Space Observatory, due for launch in 1993, before planning the next ESA move in infrared astronomy. A large passively cooled infrared space telescope, Davies says, is unlikely to reach orbit before 2005.

Peter Aldhouse



ISO tonic: the successor to IRAS, the Infrared Space Observatory, is scheduled for launch in 1993.

Graduate training reforms

London

THE Dutch university system is preparing for a quiet revolution. Starting later this year, the universities are to set up between 50 and 100 research schools, where Dutch research students will work towards their doctorates.

The plan is the latest in a series of initiatives taken in European Communities (EC) countries to improve the quality of PhD training. Following the French and the Germans, the Dutch research schools proposal is similar to the standard US graduate school model, where PhD students do research, but also take courses as part of their training.

The parallel reforms in the Netherlands, France and Germany will also have an important side effect. As the EC states move towards greater economic and political unity, the movement of research students between EC countries will become more common, and this movement will be easier between countries with broadly equivalent systems for research training. The Dutch Ministry of Education and Science realizes this, and has invited senior French, German and Belgian academics to join a committee to advise the Netherlands on its plans for reform. The committee meets this week to finalize its report, which will be published later this summer.

But the Dutch desire to develop a system of postgraduate training in consultation with other EC states finds no strong echo in Britain. British officials argue that the British system (where research students typically work in universities, attached to academic supervisors, but do not take formal courses) is in no need of reform. British research students usually gain their doctorates at a

younger age than those elsewhere in Europe, they say, with many completing their theses within three or four years.

Exactly how the Dutch research schools system will unfold is still unclear, but the change will not be imposed from above by government. Instead, the universities are expected to produce their own plans to establish research schools. The idea is that there will be several research schools for each university, with each school concentrating on a particular subject area. In some subjects, where universities have complementary research interests, two or more universities will probably collaborate to set up a research school. Each research school is expected to house about 50 research students, who are expected to complete their doctorates within four years. (In the Netherlands, these research students are employees of a university, on a modest salary.)

The Dutch reforms are not without controversy. The Netherlands Organization for Scientific Research will have a budget of 12.5 million guilders (about £3.7 million) a year to distribute among the highest quality research schools, but this is only half the figure requested by a government-appointed committee last year. Alexander Rinnooy Kan, president of the Federation of Netherlands Industry and chairman of that advisory committee, believes that the Dutch government will have to provide more money, if the research schools plan is to be "a high quality exercise".

In Germany, about 100 graduate colleges have been set up, funded jointly by the Deutsche Forschungsgemeinschaft (DFG) and the Länder, and many more are in the

pipeline. Bruno Zimmermann, from the DFG, says that ten per cent of German research students may be housed in the graduate colleges in a year's time. As in the proposed Dutch research schools, each German graduate college concentrates on a particular research theme, and provides both research supervision and taught courses. A similar network of doctoral schools is also being set up in France. Thirty-two schools were set up in the Paris region last year, and the scheme is now being extended to the rest of the country.

Britain, however, shows no sign of following the lead of its EC partners in concentrating research training into graduate schools. The Advisory Board for the Research Councils is planning to review British PhD training this year, and may compare the situation in Britain with that elsewhere in Europe. But Advisory Board officials believe there is no strong need for reform in Britain. The moves to formalize PhD training in continental Europe were influenced by a desire to lower the age at which students complete their doctorates.

Zimmermann says that British research institutes "are still very attractive to young students coming from abroad, particularly from Germany". But there is no room for complacency in the long term. Zimmermann believes that British science policymakers must eventually consider offering similar training opportunities to those being planned elsewhere in Europe. Otherwise, Britain may be sidelined as the movement of young researchers around the EC increases. Rinnooy Kan agrees. "At some point the British will be forced to adapt to the European norm."

The graduate school concept does have at least one influential supporter in Britain. Sir Mark Richmond, chairman of the Science and Engineering Research Council, would like to see each British university house a central graduate school, to accommodate all of its PhD students. Ideally, he says, these schools should offer taught courses as well as supervising research, although the additional cost may limit graduate schools to research supervision.

Richmond says his proposal is primarily a device to protect spending on research studentships in the universities from being eroded by the growing demands on the universities' budgets, rather than a desire for the British system to conform to a standard 'Euro-PhD' training — although he believes that the status of the British PhD in Europe must be looked at.

The costs of doing research are rising, says Richmond, and academic staff will be made to devote more time to undergraduate teaching, as government plans to increase the number of students entering higher education takes effect. But PhD training could be protected from these pressures by setting up separately funded graduate schools within the universities.

Peter Aldhous

SCIENCE BUDGET

Tough times extend even to the top

London

IN his maiden speech in the House of Lords last week, Lord Porter, former president of the Royal Society, described how even Nobel laureates are not immune from the funding squeeze afflicting British science.

Porter, who shared the 1967 chemistry prize for his work on photosynthesis, said his Imperial College research group has applied each year for the past four years to the Science and Engineering Research Council (SERC) for funds to buy a copper vapour laser, needed to study rapidly occurring chemical reactions. Each year the application was turned down even though on each occasion the application had been rated highly for scientific content.

Porter was speaking in a debate called to discuss the report from the Lords Science and Technology Committee on the financial crisis now facing SERC (see *Nature* 350, 264; 28 March 1991). Lord Kirkwood, a Liberal Democrat peer, contrasted SERC's predicted shortfall for 1991–92,

£40 million, with the £300 million, needed to bring the British spending per capita on science in line with that in France, Germany, the Netherlands and the United States.

The government announced last week that SERC should receive an additional £15.8 million in 1992–93. This is part of a sum of more than £30 million, which had remained unallocated when the research councils were given planning figures for their 1992–93 budgets earlier this year. But Lord Flowers, chairman of the House of Lords Science and Technology Committee, attacked this as merely the announcement of planned spending already agreed by the government. "It does nothing to reduce the present crisis in their [the research councils] funding," he said, and repeated his committee's request for an immediate emergency payment to SERC of £12 million, to lessen cuts in research grants and studentships.

Peter Aldhous

Penguins crowded out?

SIR — In your review of scientific research in Antarctica you report Wayne Trivelpiece as saying (*Nature* 350, 294; 1991) that he observed a 10–20 per cent decline in Adélie and chinstrap penguin populations near the Polish research base Arctowski in Admiralty Bay, King George Island, South Shetlands, over the past three years. He speculated that this was due to overfishing of krill in that area. He suggested that penguins be used as a monitor for krill, and that by counting penguins the krill population would be assessed. We would like to suggest that Antarctic penguins are also sensitive indicators of human interference, and that the decline in the penguin populations at Admiralty Bay may be attributed to this fact alone.

Although seemingly unconcerned, Adélie penguins react strongly to human interference during the breeding season. Heart rate is a good indicator of stress, and we found heart rates to increase by almost 50 per cent when breeding Adélie penguins were approached by a human and to increase by 270 per cent when the birds were caught and weighed in a bag^{1,2}. Adélie penguins brooding large chicks fled only when approached closer than 6 metres, but a solitary human at a distance of 20 metres from commuting penguins on a well-used walkway caused the birds to deviate by 70 metres

(ref. 2). Contrary to the view expressed in *Nature* (350, 291; 1991), we believe that tourism does adversely affect breeding penguins, almost irrespective of how “well-behaved” the tourists are.

Scientists studying penguins may also have a negative impact on the birds. Daily visits to a colony, including adult and egg measurement, reduce breeding success of the study birds when compared to non-visited colonies nearby³, and flipper bands, widely used by Trivelpiece and his co-workers, are also liable to affect birds. Kinkel⁴ reported that the use of wing tags in gulls reduced the number of birds returning to the colony, retarded the return of the others, weakened the pair bond and reduced reproductive performance. Penguins are nearly perfectly streamlined⁵ and attachment of foreign bodies including flipper bands is certain to have a negative influence on their energetics and behaviour at sea^{6,7}.

In addition, we found that aircraft operating near a base caused birds to panic at distances greater than 1,000 metres and that three days of continuous helicopter operation caused 8 per cent of the nests to be abandoned². Finally, surface-active agents such as oil, faeces and detergents originating from ships and bases destroy the waterproofing quality of the feathers and cause loss of buoyancy and insulation. Most oiled penguins die in the water, unnoticed by scientists ashore⁸.

It is therefore not surprising that penguin populations are reported to be detrimentally affected by the presence of humans. At the joint US–NZ base at Cape Hallett, Antarctica, Adélie penguins declined from 62,900 pairs in 1959 to 37,000 pairs in 1968⁹. The station was abandoned in 1973, and by 1981 the number of breeding pairs had increased to 66,000⁹. The sharp decline in penguin numbers at Cape Royds between 1955 and 1963 has been attributed¹⁰ to interference by visitors on foot and to helicopters flying over the colony. Since the restriction of human activity in that area, the population has recovered. Recently, Woehler *et al.*¹¹ found that at Shirley Island, near Australia's Casey station, the Adélie penguin population had increased by 209 per cent everywhere but in the vicinity of the base, where numbers were stagnating.

Scientists in our department have spent several field seasons in Admiralty Bay, and report that the area is increasingly affected by tourism. Tourists and personnel often approach and enter colonies within the ‘sites of special scientific interest’. And the operation of scientific vessels and supply ships, often with helicopters, further increases disturbance throughout the breeding season. King George Island is the most densely populated area in Antarctica, with eight stations being operated year round and a large number of summer camps, with all the

associated problems of disturbance, sewage and pollution.

If penguin numbers are to be used to monitor the krill population, the study site as well as the methods employed by the scientists in the field need to be very carefully selected. Neither is the case in the studies conducted on King George Island.

BORIS CULIK
RORY WILSON

*Institut für Meereskunde
an der Universität Kiel,
Düsternbrooker Weg 20,
D 2300 Kiel, 1, Germany*

1. Culik, B., Adelung, D. & Woakes, A. J. in *Antarctic Ecosystems. Ecological Change and Conservation*. (eds Kerry, K. R. & Hempel, G.) 177–182, (Springer, Berlin, 1990).
2. Wilson, R. P. *et al. Polar Biol.* (in the press).
3. Lishman, G. S. *Ibis* 127, 84–99 (1985).
4. Kinkel, L. K. *Auk* 106, 619–624 (1989).
5. Nachtigall, W. & Bilo, D. J. *comp. Physiol. A* 137, 17–26 (1980).
6. Wilson, R. P. *et al. Polar Biol.* 10, 77–80 (1989).
7. Culik, B. & Wilson R. P. *J. exp. Biol.* (in the press).
8. Culik, B. *et al. Mar. Pollut. Bull.* (in the press).
9. Wilson, K.-J., Taylor, R. H. & Barton, K. J. in *Antarctic Ecosystems. Ecological Change and Conservation*. (eds Kerry, K. R. & Hempel G.) 183–190 (Springer, Berlin, 1990).
10. Thompson, R. B. in *Adaptations within Antarctic Ecosystems*. (ed. Llano, G. A.) 1177–1180 (Smithsonian Institution, Washington, DC, 1977).
11. Woehler, E. J. *et al. Mar. Ornithology* (in the press).

Secret Service

SIR — As the Office of Scientific Integrity (OSI) draft report of the investigation of the Weaver *et al.* 1986 *Cell* paper remains confidential, I will not discuss details of its contents. But I should like to correct two significant errors in your article “Secret Service as ultimate referee” (*Nature* 350, 553; 1991).

(1) The Secret Service possesses extraordinary expertise in many areas, notably in forensic work, and they contributed very significantly to the OSI investigation, but they were not involved in the statistical analyses to which you refer, neither is it anywhere represented that this is the case. You apparently have received incorrect information or made an unwarranted inference.

(2) Contrary to your assertions, the date of creation of the subcloning data is not known with certainty, and is not stated in the draft investigative report. To be consistent and in sequence with related experiments, the subcloning data would have had to be created around June 1985.

The substantive issues you raise about the statistical analyses are important, and we look forward to further discussion of these matters at the appropriate time.

SUZANNE W. HADLEY

*Office of Scientific Integrity,
National Institutes of Health,
Room 207, Bldg 1,
Bethesda, Maryland 20892, USA*

■ See opposite page.

Honest blunders

SIR — These days, we hear a lot about fraud and deceit in science. What we collect, however, is the honest blunder. When Einstein was studying to be admitted to the Zurich technical university, he was doing rather badly: he scored mainly 1s (on a scale from 1 to 6). But in the second half of the year things suddenly went better, and Einstein's marks went up to 6. According to a report of a Canadian historian¹, this was obviously due to the fact that the school opened a brand-new physics laboratory, in which Einstein could do his own experiments.

Only after he had published this hypothesis was the historian told² that, in the same year, the school had reversed its scale, so that it now ran from 6 to 1.

That is the kind of blunder we like.

We would appreciate hearing about (preferably your own) most embarrassing moments in science. We want to expand our collection and, with your help, perhaps blunder into publishing a book on the subject.

HANS VAN MAANEN

*Het Parool,
PO Box 433, 1000 AK Amsterdam,
The Netherlands*

1. Pyenson, L.: *The young Einstein*. p.12. (Hilger, Bristol, 1985).
2. Stachel, J. J. (ed.): *The collected papers of Albert Einstein* I. p. 14, note 3. (Princeton University Press, 1987).

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

Baltimore declares O'Toole mistaken

Dr Margot O'Toole's comment on the OSI draft report reproduced in *Nature* two weeks ago, has drawn the following reply from Dr David Baltimore, one of the authors of an article alleged to include fraudulent data.

DOCTOR Margot O'Toole's recent comments on the draft report of the National Institutes of Health Office of Scientific Integrity (*Nature* 351, 180; 16 May 1991) create a misleading impression and therefore require a response. The issues she raises have all previously been answered, often several times, but because many are not familiar with the details, I feel that it is necessary to demonstrate publicly that her charges lack substance.

I have already submitted comments on the draft report of the investigation conducted by the OSI into the controversy over the *Cell* paper. In them, I acknowledged certain errors of judgement on my part, commended Dr Margot O'Toole for her courage and insight and volunteered to participate in the continuing dialogue with Congress and other government bodies responsible for the oversight of scientific research supported by public funds.

The OSI report contained no allegations or findings that I participated in any falsification or fabrication of data, but its criticisms of me focused upon my defence of my co-author, Dr Thereza Imanishi-Kari, particularly at the May 1989 hearings before the congressional subcommittee. My comments explained that my defence of her grew out of the trust and respect I had for my collaborator's demonstrated abilities as a scientist, my understanding that the molecular analysis conducted in my laboratory validated Dr Imanishi-Kari's conclusions and my belief in the efficacy of the peer review process.

Since then, Dr O'Toole has commented publicly about my statement, and released a lengthy response to the OSI draft report that enumerates the charges she is now making. These are the remarks published in *Nature*.

Dr O'Toole's response, unlike the OSI draft report, contains allegations that I was aware before the completion of the OSI report that statements in the *Cell* paper were untruthful, that I was remiss in submitting a letter of correction to *Cell* citing the unpublished data by Dr Imanishi-Kari which Dr O'Toole states were false and that I have publicly attacked Dr O'Toole's competence and motives. For those reasons, I believe it is incumbent upon me to clarify these matters.

Dr O'Toole's response contains new charges that are different from her original constructive questions on matters of science, and also includes certain overstatements and errors. Any assessment of her claims must take into consideration a number of facts she has not mentioned, and I have therefore provided the information that follows.

■ The importance of the notebooks discovered by Dr O'Toole.

Dr O'Toole begins her recitation of the history of the controversy by stating that in May 1986, she came across laboratory records "for experiments on which the central claim was based" and that it was "obvious" from those records that the experiment had not yielded the published results. She goes on to state that this flaw "remains evident", and that Herman Eisen of Massachusetts Institute of Technology (MIT) said to her that the evidence "appeared to indicate fraud".

The 17 pages of data Dr O'Toole copied were taken not from the notebooks of Dr Imanishi-Kari, but from those of Dr Imanishi-Kari's postdoctoral fellow, Moema Reis. Dr Reis's data have not been challenged. The

The article based on Dr Margot O'Toole's comment on the draft report by the NIH Office of Scientific Integrity (*Nature* 351, 180; 16 May 1991) has understandably provoked controversy. At issue is the authenticity of the data on which an article published in 1986 is based (Weaver, D. et al. *Cell*). This symposium of documents includes replies to O'Toole from two of those whom she criticized — Dr David Baltimore (president of the Rockefeller University) and Dr Herman N. Eisen (MIT) — as well as a much-abridged version of Dr Thereza Imanishi-Kari's comment on the OSI draft report, published in March (*Nature* 350, 263; 28 March 1991). Baltimore's own response to the same report appeared on 9 May (*Nature* 351, 94; 1991). Others who have signalled their wish to join this discussion will be accommodated as space permits.

major element of the 17 pages that appeared to contradict the study were data supposedly derived from a control mouse. These differed from the published data on control mice.

It was discovered quite quickly, at the first meetings in May 1986 when Dr O'Toole raised her challenge, that the mouse in the 17 pages had been mistyped — it was not a control mouse and was in fact, transgenic. Data on truly normal mice were generated, and it was those findings that were used in the paper.

Although Dr O'Toole has acknowledged these facts in the past, she continues to charge that the 17 pages contradict the published paper, while neglecting to mention that the pages have proved to be irrelevant. By citing Dr Eisen's comment, she leaves the impression that the data she found are evidence of fraud. This issue has been defini-

tively resolved and therefore, Dr O'Toole's statement that the discrepancy "remains evident" is incorrect, and her reference to the data she discovered and Dr Eisen's reaction to it is misleading.

■ **Dr O'Toole's challenge to Figure 1.** Dr O'Toole states that the paper's Figure 1 is "not truthful". She repeats at several points that experiments "central to the paper" had been performed, but did "not yield the published results". Dr O'Toole does not identify the experiments involved, but she has made such statements before when discussing the issues surrounding the reagent BET-1. Her original challenge focused largely on her belief that BET-1, a reagent used to distinguish between the antibodies that were the object of the study, and described in Figure 1, did not work. The airing of her views in 1986 led quickly to a realization that the specificity of BET-1 had been somewhat overstated in the paper.

Dr O'Toole repeats the allegations although this mistake was subject of a letter of correction to *Cell* in October 1988, two and a half years ago. This overstatement has been deemed a mistake by the Tufts, MIT and NIH reviewers and not the product of fraud.

Dr O'Toole persists in challenging Figure 1. Every academic committee or government agency that has reviewed the matter has rejected her contention that BET-1 does not work at all. Indeed, the January 1989 report of the first NIH panel pointed out that data showing the effectiveness of BET-1 could be found in Dr O'Toole's own notebooks.

■ The subclone data submitted with the May 1989 letter of correction to *Cell*.

One of Dr O'Toole's most serious charges is that I published information as part of the 1989 letter of correction to *Cell* that I knew was false. Dr O'Toole states that the experiments on subcloning of the hybridoma wells in Table 2, described in the letter, were not in fact performed. Although she does not expressly state that I knew the data to be fabricated, her complaint that I submitted the letter to *Cell*, even after she told NIH in November 1988 that what have become known as the "June subcloning" experiments had not been performed, leaves the impression that I made a knowing misrepresentation. This conclusion is untrue.

First, I wish to state categorically that I have not throughout the history of this matter made a statement that was known to me to be untrue, or which I even suspected was untrue. In addition, I have never heard from Dr Imanishi-Kari that she did not perform any of the experiments described in the *Cell*

paper, or in the various published corrections. The fact that, in 1987, I called for a full review of the matter by NIH indicates that, in my mind, we had nothing to hide.

In November 1988, the first NIH panel to study the matter issued its draft report. Dr O'Toole responded and charged for the first time that the panel was drawing its conclusions from experiments which she said Dr Imanishi-Kari told her had not been done. She stated that Dr Imanishi-Kari told her during the course of the review by the Tufts *ad hoc* committee that no subcloning analysis of the Table 2 hybridomas had been performed and Dr O'Toole denied that subcloning data had been reviewed during those meetings. As I was not present during the Tufts meetings, and I had never heard Dr Imanishi-Kari make such a statement, I had no personal knowledge with which to verify or refute Dr O'Toole's remarks.

In December 1988, NIH solicited more information from Dr Imanishi-Kari and from Dr Henry Wortis, Dr Brigitte Huber of Tufts and Dr Robert Woodland of the University of Massachusetts, who had conducted the review for Tufts. All confirmed that in May 1986, they reviewed the data in support of Table 2, that the data supported Table 2 and its conclusions, and that the subcloning analysis was discussed with Dr Imanishi-Kari. The reviewers met with Dr Imanishi-Kari on 16 May 1986, when Dr O'Toole was not present.

The NIH panel stated in its final report that it carefully reviewed Dr O'Toole's letters, but found that no modification to its report was necessary. I believe, then, that my acceptance of the data as authentic at the time was justified. No forensic analysis had as yet been performed. Moreover, the OSI draft report states that Dr Reis told the investigators that she participated in performing the experiments.

■ **The isotyping experiments.** Dr O'Toole states that Dr Imanishi-Kari told me and the others assembled for the meeting with Herman Eisen of MIT in June 1986 that, in preparing Table 2 for publication, "she relied completely on her prior expectation of what the results would be", and she states that Dr Imanishi-Kari admitted in my presence "that a large series of the published experiments had not even been performed."

Dr O'Toole's failure to specify which experiments she means complicates any effort to respond, but the questions concerning the isotyping experiments on the Table 2 hybridomas were answered long ago. The experiments were in fact done, but — unknown to me at the time — the exact materials used were misidentified in the paper. The point was dealt with in the letter of correction to *Cell* in November 1988, and it is unclear to me why this issue would be raised now.

■ **The impact of the Secret Service briefing in 1989.** Dr O'Toole faults me for defending Dr Imanishi-Kari in my May 1989 testimony before Congress, even though the

Secret Service and members of the staff of the congressional subcommittee had met me before the hearings to review the forensic findings. At that meeting, the Secret Service's work was still in progress; some of the irregularities the agents had uncovered in the notebooks were described to me in a very fragmentary and unsystematic fashion, without the aid of demonstrative charts or formal reports. The briefing did not touch on the evidence stressed in the recent OSI draft report — the analysis of paper used in the printer attached to the gamma counter.

The types of difficulties generally described to me — for instance, the obvious changing or adding of dates — seemed at the time to be as consistent with Dr Imanishi-Kari's known reorganization of her original materials as with any effort to commit fraud. When I met the Secret Service, I discussed at length how the notebooks had come to be put together.

For the reasons cited above, and because I had a high regard for Dr Imanishi-Kari, I did not abandon my faith in her at the time. In hindsight, as I have explained in my comments to OSI, it would have been prudent to step back and review the evidence before commenting further; my failure to do so demonstrates only my high degree of trust in a fellow scientist, and not any intent to deflect or obstruct an investigation.

Dr O'Toole expresses her view that I was not forthcoming with Congress because I did not explain that the pages of original data for the paper were organized into book form before being submitted to NIH in July 1988. I have made no secret of the fact that, when I first saw the records, they were not in books, but were a mixture of notebooks, loose paper and data sheets. Dr Imanishi-Kari acted on the advice of a lawyer and against my advice in pulling them together into an organized form to facilitate their review by scientists of NIH. I have stated this on numerous occasions over the years, as well as in interviews with OSI and the Secret Service.

■ **The 1991 retraction.** When summarizing the history of the matter, Dr O'Toole refers to the issuing of the OSI draft report and states: "Only then, five years after he learned of the problems, did Dr Baltimore retract the paper." This sweeping statement conveys the impression that "the problems" contained in the OSI report were known to me for five years, but the factual circumstances are quite different.

The serious allegations contained in the OSI report first became known to me in March this year, when the report was sent to me for comment. Five years ago, "the problems" identified by Dr O'Toole were of an entirely different nature. Her written memorandum to Dr Eisen of MIT in June 1986 specified the problems as she saw them. Then, Dr O'Toole expressed her view that the data did not support the paper's conclusion, and she proposed alternative explanations for the observed phenomena. In 1986, Dr O'Toole challenged the specificity of

BET-1 and claimed that Table 2 understated the frequency of certain findings with respect to normal mice.

She challenged the paper's finding that it was the endogenous gene, and not the transgene, that was being expressed when idiotype-positive antibodies — that is, those characteristic of the transgene — were found in hybridomas from transgenic mice. Her argument was that the assays used were inadequate and insufficiently sensitive to detect the presence of the transgene. In other words, she questioned the accuracy of the paper's conclusion on matters of science.

In 1986 and 1987, Dr O'Toole was emphatic that she was alleging error, and not charging fraud. When she appeared before Congress in April 1988, she continued to stress that she was concerned only about error and basic scientific principles. Her allegations about the fabrication of the June subcloning data were not raised until late November 1988. Thus, her suggestion that I ignored charges of the seriousness of those contained in the OSI draft report for five years is not accurate.

Dr O'Toole also states that I "consistently and falsely" maintained that her objections were no more than alternative interpretations of valid data. The Tufts committee, Dr Eisen of MIT, and NIH reached precisely this conclusion and Dr O'Toole herself insisted in 1986, 1987 and 1988 that her only concern was for the validity of the science.

■ **Compliance with NIH recommendations.** Dr O'Toole states that the authors did not comply with the recommendation of the NIH panel contained in its January 1989 report. This statement is incorrect. We complied with all of the recommendations and the director of NIH accepted our response as appropriate.

■ **The meeting of Dr O'Toole, the authors and Dr Eisen of MIT.** Dr O'Toole provides a detailed description of the June 1986 meeting with Dr Eisen of MIT. Her account is couched in terms that incorrectly suggest that we were told then of a series of misrepresentations and that at the time I advised against a further public airing of her views.

First, Dr O'Toole states that I "acknowledged" at the time that the published results could not be based upon the data she brought to the meeting, that is, the 17 pages from Dr Reis's notebook described above. The 17 pages, as I have explained, had already been determined to be irrelevant.

Second, Dr O'Toole states that Dr Imanishi-Kari "admitted that a large series of the published experiments had not even been performed". It is simply not true that any large series of experiments had not been performed. Perhaps this overly broad reference is to the fact that the Table 2 idiotype type experiments were not done on the particular hybridomas indicated, but were performed on other hybridomas. That has already been the subject of a published correction.

If Dr Imanishi-Kari had stated at the meeting that she did not perform any of the

experiments, I would have immediately questioned the veracity of the paper. My remembrance that she made no such statement is supported by the sworn testimony of Dr Eisen before Congress that the question of tests not being performed never came up at the meeting.

Finally, Dr O'Toole states that I told her I would personally oppose any effort she might make to get the paper corrected. I do recall Dr O'Toole discussing, just as the meeting was ending, the prospect of her sending a letter to *Cell* specifying her criticisms. I did say, I believe, that the authors would then probably respond to her letter. I believe that such a response would have been appropriate in normal scientific disagreements.

Dr O'Toole's description of this exchange differs significantly from the account she submitted to Congress in April 1988 as part of her written chronology. There she states that I proposed that she could write to the journal and noted that I would respond. She goes on to say that she then "stated that [she] consider[ed] [her] responsibilities discharged and intended to drop the matter." Dr O'Toole's own previous version of the events contradicts the theory she is advancing now: that she wished to pursue the matter further in the journal, but that I intimidated her from doing so.

■ **My views on fraud in science.** Dr O'Toole purports to describe my views and incorrectly states that I maintain that "false claims do not have to be corrected because other investigators will stumble on error and clear things up eventually." This is a gross parody of my stated belief that the scientific process is the best means to test the scientific validity of published claims. As I have repeated in my articles on the subject, my congressional testimony and, most recently, in my response to the OSI draft report, consciously false claims, or fraud, by a scientist can never be excused or condoned.

■ **Dr O'Toole's claim that she has been personally attacked.** Dr O'Toole states that throughout the history of this matter, there have been "attacks on [her] competence and motives". She also states that she was subjected to "five years of slander and libel from Drs Baltimore, Eisen and Imanishi-Kari." Dr O'Toole makes this accusation, yet she cites not one example of any comment by me in which I publicly disparaged her or her ideas.

At the June 1986 meeting, I stated that I considered Dr O'Toole's criticisms to be imaginative and thoughtful. When Dr Eisen wrote a report about the meeting, he remarked, "I do not think that I or anyone present at the meeting felt that Margot O'Toole's disagreements were frivolous." None of my written statements or testimony about the matter raised any question about Dr O'Toole's abilities as a scientist or her motives. I told the congressional subcommittee that Dr O'Toole's criticisms "were a rational and appropriate part of the scientific

process". My comments to OSI stated that her analyses were insightful, her expressions of concern were proper and appropriate and her motives were pure.

While Dr O'Toole has now directly attacked my honesty and integrity, none of my previous remarks nor any of the remarks in this statement were intended to criticize her personally, impugn her abilities as a

scientist or question her motives. Rather, I have made this statement in the hope that any assessment of the validity of her comments will be a measured one, based upon a consideration of all of the facts and the entire record of this controversy, including Dr O'Toole's own previous statements on the matter.

David Baltimore

Origins of MIT inquiry

Dr Herman N Eisen, who conducted the inquiry at MIT in 1986, denies Dr O'Toole's account of its proceedings, and other charges.

DOCTOR Margot O'Toole makes a series of assertions that effectively charge me and many others with dishonest and irresponsible behaviour. These assertions cannot go unchallenged. In what follows, I address several of her more extreme statements, those that from personal knowledge I know to be inaccurate or grossly to misrepresent the true events.

It is important to clarify the circumstances surrounding the initial inquiry at MIT, now characterized by Dr O'Toole as a "cover-up". I met and spoke to Dr O'Toole on three occasions, all in the spring of 1986. First, she told me in March or April of experimental and personal difficulties she was experiencing in Dr Imanishi-Kari's laboratory. Second, on 30 or 31 May, she visited me at Woods Hole to describe her concerns about the validity of the Weaver *et al.* paper. Third, on 16 June 1986 she and I met the paper's authors (Baltimore, Imanishi-Kari and Weaver) to consider the memorandum she had prepared at my request.

In that memorandum (dated 6 June), Dr O'Toole elaborated clearly and at length on what she saw as four principal sources of error in the paper. However, it contains no suggestion that reported results were based on nonexistent or fraudulent data. I have read that memo many times since, searching for indications of the fraud that Dr O'Toole later proclaimed so vigorously but I have found none.

I have, therefore, long been puzzled by Dr O'Toole's turn-around in asserting — beginning, I think, two years later — that fraud was evident at the very start. Recently, however, I have come upon a hint of an explanation: in testimony before the Dingell subcommittee on 12 April 1988, Dr O'Toole stated that her memo of 6 June 1986 had been edited before being submitted to me to remove "any language that might possibly imply that she was alleging any misconduct".

She presumably took this action to avoid making an allegation of fraud without substantial evidence, as such a charge might expose her to legal action. While her trepidation is understandable, the result was a carefully manicured memorandum that contained no hint of fraud. In consequence, I

was presented with a document that, except for its intensity and length, had all the hallmarks of a typical scientific dispute, the kind editors see repeatedly between authors and an intelligent, intensely engaged and critical reviewer of a manuscript.

Thus it is not inappropriate to ask: who misled whom? It is ironic and sad that, instead of recognizing that she bears some responsibility for creating a misleading situation, Dr O'Toole now characterizes the initial inquiry at MIT as a "cover-up". Given her choice of words, I also find it remarkable that those of us who were involved in the inquiry are accused of slander and libel.

I now consider the particular statements of Dr O'Toole's that are especially serious and that I find to be completely at odds with the events as I know them.

■ Dr O'Toole asserts that, at the 16 June 1986 meeting, "Dr Imanishi-Kari again admitted that a large series of the published experiments had not even been performed and that some which were performed had not yielded the claimed results." I deny that such a statement, amounting to a clear admission of fraud, was made at that meeting in my presence — and I was present throughout. Had such a striking confession been made at the time, it is difficult to understand why Dr O'Toole failed to mention it in her testimony before the Dingell subcommittee in 1988.

■ Dr O'Toole asserts repeatedly that the 17 pages she obtained from Dr Reis's notebook provided obvious evidence of problems with the paper. Her allegation grossly oversimplifies the complexity of the material. Walter Stewart, testifying two years later before the Dingell subcommittee, stated that the language in these pages was in "Portuguese or some other language" and that the pages were "extremely confusing because they consisted mainly of numbers". Earlier, I had asked him how long it has taken him to make sense of the 17 pages. He replied in one word: "weeks".

At the time, however, the presence in those pages of some data that did not agree with published material did not seem unusual. Notebooks commonly contain much imperfect data accumulated on the way to obtaining definitive results for publi-

cation. As it was not made clear that those pages may have represented the only data for one of the tables in the paper, it is unreasonable to contend, as Dr O'Toole now does, that a brief examination of them would have revealed that they were *prima facie* evidence for serious problems. Indeed, if her contention were correct, I find it remarkable that she failed to identify the 17 pages in the 6 June memo which formed the cornerstone of our inquiry.

■ Dr O'Toole now asserts that, during the 16 June meeting, Dr Imanishi-Kari admitted that Figure 1 in the *Cell* paper was not a truthful rendering of experimental results. This version is incorrect. The admission made concerned the exaggerated statement in the text of the paper indicating that the BET-1 antibody was absolutely specific (exhibiting no cross-reactivity). All at the meeting agreed that this exaggerated statement was an error and there was considerable discussion then and later about sending a note of correction of *Cell*. I regret that I did not more forcefully urge the authors to do so.

■ Dr O'Toole asserts that I have slandered and libelled her over the past five years. This is not true. I fear she confuses disagreement with slander and libel. On the contrary, I have always indicated that her analysis of the serological part of the disputed paper was cogent and thoughtful. But I did not agree with several of her arguments. For example, her assertion that some of the results could be explained by heterodimer formation, which I take to mean μ - γ chimaeric molecules, is highly implausible. And her assertion that transgene products could have been overlooked because sensitivity of radioimmunoassays for detecting immunoglobulin idiotype is much greater than mRNA assays is, I believe, incorrect. On the other hand, I agreed with her emphasis on the problematic specificity of BET-1, though her statements in this regard are probably exaggerated. It is unfortunate that she mistakes such disagreements for slander and libel.

■ Dr O'Toole asserts that "I lost my career because I would not go along with Dr Imanishi-Kari's pressure to misrepresent my own results. I informed Dr Eisen of this pressure and he had a direct responsibility to help me. He failed in this and other responsibilities to me."

I deny that Dr O'Toole ever related such events to me. This is the first I have heard of this awful charge. Indeed, Dr O'Toole gave an entirely different version of what she reported to me in her testimony to the Dingell subcommittee on 4 April 1988.

■ Dr O'Toole implies that her career has suffered because I refused to help her. In fact, she never asked for my help. At the end of the May 1989 Dingell subcommittee hearings, Dr John Deutch, then provost of MIT, and I agreed with Mr. Dingell that we would try to help Dr O'Toole find a job. Although she may have subsequently contacted others at MIT, she never communicated with me or sought my help.

■ Dr O'Toole makes an obscure assertion

that some critical data that helped unravel the "tangled web" lay buried in NIH files in "Dr Eisen's grant application", as though I were somehow responsible for some unsavoury data. The application was a Program Project involving five independent investigators (including Imanishi-Kari and myself as principal investigator). In such applications, each investigator is responsible for his or her own data; the data in question would have appeared only in Dr Imanishi-Kari's portion.

■ It has been asserted that Dr O'Toole lost her job at MIT because she challenged the Weaver *et al.* paper. She now says that "this charge is essentially true, but the facts were a little complicated". In fact, they are not. Dr O'Toole was appointed a postdoctoral fellow in June 1985 with the understanding the position would terminate on 31 May 1986. She was ineligible for more than one year of

support from the NIH training grant that provided her stipend on account of previous years of support. Moreover, it was known in 1985 that Dr Imanishi-Kari was due to move from MIT to Tufts in June 1986.

Looking back to the beginning of this tragic affair, it is hard to imagine less favourable circumstances for reaching a judicious resolution of the matter. The departure of the two principals (Dr O'Toole two weeks before Dr Imanishi-Kari) as it began left little room for the repeated informal personal interactions that might have helped clarify the scientific issues and the ambiguities surrounding Dr O'Toole's charges.

Nobody disputes Margot O'Toole's stated belief in the importance of integrity in science. She is to be admired for her tenacity, but her attacks on those who disagree with her are unwarranted. **Herman N. Eisen**

OSI's conclusions wrong

What follows is a much-abridged version of Dr Thereza Imanishi-Kari's formal response to the OSI draft report published at the end of March.

THE Secret Service evidence fails to support the conclusions of the draft report. This assertion comes from my analysis of the draft, the summaries produced by the Secret Service and the published testimony of the Secret Service before the House Committee. My analysis was conducted without access to the full Secret Service reports, to any chromatographic or other data produced by the Secret Service, to any of the original documents examined or even an index of the full record of the Secret Service examination.

Most outrageously, until the draft report was issued, my defence was conducted without knowledge of the specific accusations against me. I cannot disprove any of the Secret Service allegations without access to the original laboratory notebooks and tapes. Yet I have been repeatedly denied access to this material by the Office of Scientific Integrity (OSI).

This unfair denial of access should be seen in the light of the statement by Dr Suzanne Hadley, then acting director of OSI, (quoted in the *New York Times*, 3 May 1991) that "If we showed a scientist all the evidence . . . before we wrote the draft report, we would run the risk of giving him information that he could then use to cover his tracks".

Dr Hadley fails to state the obvious, that under these unique rules, the accused is denied the means to establish innocence.

The draft report places major emphasis on the evidence of forensic and statistical analysis. This is the weakest of all possible forms of evidence, and extrapolated conclusions from flimsy experimental results. In contrast, the draft ignores all the direct statements made by me, Dr Reis, Dr Weaver, Dr Baltimore, Dr Wortis, Dr Huber, Dr Woodland, Dr Eisen or by any of the several laboratory

investigators who worked at Massachusetts Institute of Technology (MIT) at the time the research was done.

This evidence, which I find to be very supportive to my argument, is dismissed out of hand. It is as if the author of the draft report believes all of these scientists are part of some giant conspiracy and the only person whose testimony has credence is Margot O'Toole.

The draft and Dr O'Toole suggest that I wrote new data after allegations surfaced. In fact, all I did was to take pages already written and organize them in a coherent fashion. I did not write new data.

The experiments known as "the June subcloning" consist of radioimmunoassays (RIAs) of the antibodies produced by cloned hybridomas. These hybridomas were cloned from wells described in Table 2 of the *Cell* paper. The results of these RIAs were not published then, but were reviewed in May 1986 by Drs B. Huber, R. Woodland and H. Wortis and again by the first National Institutes of Health (NIH) Panel in September 1988.

In November 1988, in response to the report of the first NIH Panel, Dr O'Toole alleged that she had been told such experiments did not exist.

The draft report now says: "OSI believes that several significant aspects of the June subcloning data render them highly suspect, especially in light of Dr O'Toole's assertion that she was told that no subcloning experiments had been done."

Dr O'Toole's current claim that I denied cloning the Table 2 wells stems from discussion at the May 1986 meeting between me, O'Toole, Huber and Wortis. In her response to the first NIH report she wrote: "At the ▶

meeting with Drs Huber and Wortis, Dr Imanishi-Kari stated emphatically that, besides tests recorded in the 17 pages already in my possession, no other test had been done on the Table 2 cells."

This statement is accurate, but it means that no further serological tests were performed on supernatants from Table 2 cells. Indeed, I did say . . . that no isotyping was performed on those supernatants. This, however, is very different from the assertion Dr O'Toole makes "that Dr Imanishi-Kari told me in May 1986 that no subcloning analysis of Table 2 hybridomas had been done."

For me to have told O'Toole, Huber and Wortis in May 1986 that no cloning was done would have been ludicrous, as there was a record of this cloning in Dr Reis's R-1 notebook. This is a clear unchallenged record of the cloning of Table 2 hybridomas. This is proof that the cloning was done and directly contradicts Dr O'Toole's assertions.

If the cloning of Table 2 wells was crucial, and if I had admitted to Dr O'Toole that such cloning was not done, why does her June 1986 memo to Dr Eisen make no mention of either? If I had admitted to her at the second meeting with Dr Wortis and Dr Huber that I did not perform "the crucial" cloning and subsequent RIA experiment, why did she fail to mention this at the 16 June meeting with Eisen, Baltimore, Weaver and me?

The counter-argument might be made that the "June subcloning" data were prepared in anticipation of the visit of Huber, Woodland and Wortis. But I had no foreknowledge of the questions they would ask, and from the June memo of O'Toole to Eisen and the later memo of Wortis, it should be apparent that O'Toole's allegations were not focused on questions that would have been met by the "June subclone" analysis.

It is remarkable that OSI chooses to believe the uncorroborated assertion of Dr O'Toole over direct testimony of Drs Reis, Wortis, Huber and Woodland. The draft accepts her current version of events rather than her own written memo of June 1986.

Of the evidence on the June subcloning data, the draft report appears to view the forensic analysis as the most damning. It asserts that "forensic analyses demonstrate compellingly that the June subcloning data are not authentic, and therefore fabricated".

The forensic evidence of the Secret Service [consists of an] analysis based on colour of paper, composition of ink, intensity of ink and font of printer The author(s) of the draft report, and the scientific advisory panel, consider the crucial portion of its evidence to be chromatographic analysis of the inks used for printing the results of gamma counts obtained with a Beckman counter. In their minority opinion, Drs Hugh O. McDewitt and Ursula Storb say why they view these data as important [and add]:

"Unfortunately, the panel was not given access to the actual chromatograms which led the Secret Service to make the conclusions referred to above Ultimately, the

actual chromatograms need to be examined."

The reason for the lack of chromatographic evidence is that either none was produced or [that] the evidence failed to support the claims of the draft report.

The Secret Service and the draft report make much of the comparison of the gamma counter tapes related to the RIA of the "June subclones". However, this comparison is flawed by the fact that the gamma counter tapes [in the I-1 notebook] were compared to tapes from different types of counters. But tapes from a beta counter do not produce output that matches the format of Beckman gamma counter tapes. Such an "apple versus orange" comparison is utterly lacking in scientific significance.

What fraction of the total output of Beckman gamma counter tapes was examined by the Secret Service? If (as assumed by the minority) the Secret Service analysed the majority of counter tapes generated during the period 1981-86, where are the matching Beckman gamma counter tapes?

[There are also difficulties about the colour of the tapes.] According to the draft, these particular tapes have a green hue, and the claim is made that there were no "green" tapes in the examined laboratory record dated later than January 1984. However, tapes existed in many shades of "yellow".

Furthermore, when new rolls of paper were not available, left-over paper comprising the ends of old rolls was used. Inspection of additional counter tapes reveals sudden brief changes in tape colour that can last for only a few days. The "failure" to find other Beckman gamma counter green tapes from 1985 loses significance in this context.

Although a complete record of the tapes available to OSI is not available to us, it appears from the draft that only a small percentage of pages of annual output from the Beckman gamma counter were examined. This means that, contrary to the assumptions of the draft and the minority report, the record examined was far from the majority of tapes produced at the time. The Secret Service findings are [thus] further eroded.

The fundamental conclusion of the data that emerges from this RIA analysis is that some hybridomas produce idiotype-positive immunoglobulin yet are negative for the transgene allotype The draft argues for its conclusions as if the experiments on these pages were the unique laboratory record of evidence on this point. If these pages provided such unique evidence, there would be motive to falsify them. But . . . there are other data throughout the I-1 and R-1 notebooks which make the same point. This evidence exists on [I-1] pages unquestioned by the Secret Service and in notebooks produced by Dr Moema Reis and which are completely unquestioned.

The first allegation of fraud regarding these data was made by Dr O'Toole in a letter to OSI dated 6 November 1989, and is interesting in that it uses formulations remi-

niscient of the Secret Service report. Is it possible that Dr O'Toole made her remarkably prescient charge of fraud because she already knew the Secret Service conclusions?

[Dr Imanishi-Kari's statement includes at this stage a closely similar criticism of another important series of experiments yielding the "January fusion" data, and concludes:] Since other data establish the low frequency of the idiotype in normal mice, there was no motive for creating fraudulent data. As Margot O'Toole did not allege that this experiment never existed there was no reason to "concoct" these data. If the draft had failed to allege that the "January fusion" tapes were fraudulent, they would provide evidence that matching "green" counter tapes were produced in the laboratory as late as January 1985. If "green" counter tapes were produced in January 1985, the claim that the "June subcloning" tapes were produced in 1981-82 must be discarded. In other words, an accusation for the "January fusion" was needed to support the accusations regarding the "June subcloning".

You have no idea how upsetting it is to find that after five years of controversy and two years of OSI review, an investigation still gets the most fundamental part of the story — the question of the BET-1 reagent — all wrong. It is demoralizing to find that, after all this time, there is still such a basic misunderstanding of the reagent first questioned by Dr O'Toole five years ago.

The draft report states that "the BET-1 reagent was key to the reliability and accuracy of the results . . . [it] played a crucial role in the detection of the transgene. A failure of BET-1 to react with appropriate specificity would open the possibility that Dr Imanishi-Kari was in many cases detecting transgene, rather than endogenous, genes."

Unfortunately, this gets the BET-1 story completely backwards. BET-1 is a monoclonal reagent raised against the μ_a allotype. It recognizes the product of the transgene μ_a . A lack of specificity of BET-1 would lead to recognition of μ_a (endogenous) and therefore to an overestimation of the amount of transgene product. This means that, used as described in the *Cell* paper, any lack of BET-1 specificity would strengthen the conclusions of the authors. Therefore, I had no motive to falsify data on BET-1.

The OSI has asked repeatedly whether I performed iodination experiments. I have never studied the effects of iodination on antibody specificity and I did not start such studies in the middle of our transgene project. When asked by the OSI, I denied performing such experiments. The only reason that the draft says that I did . . . is to justify its use of the forensic data concerning certain experiments. If I had no interest in proving the specificity of BET-1, or the effects of iodination, then the forensic data pertaining to the dates would be rendered meaningless. This is a result which OSI could not tolerate.

Thereza Imanishi-Kari

Split decision on the mantle

Dion L. Heinz

DOES the composition of the Earth's lower mantle differ from that of the upper mantle, or is the mantle of uniform composition? Two new papers, by Li and Jeanloz¹ and Wood and Nell², come to differing conclusions. Both teams measure the electrical conductivity of mantle materials, and compare their results with independent measurements of the electrical conductivity of the Earth's mantle. Li and Jeanloz argue that the lower mantle is either more iron-rich than the upper mantle, or that it contains a significant component of volatile elements³, whereas Wood and Nell conclude that the lower mantle is of the same composition as the upper.

Much of the research on the composition of the lower mantle has concentrated on what is represented by a discontinuity found on seismograms at a depth of 670 km. All major upper-mantle minerals (such as olivine, pyroxene and garnet) transform to either magnesium-iron silicate perovskite ((Mg,Fe)SiO₃ in the perovskite structure), or magnesium silicate perovskite and magnesio-wüstite ((Mg,Fe)O in the rocksalt structure) at lower-mantle conditions. If the jumps in density and seismic velocity at the 670-km seismic discontinuity are due to a change in composition in addition to the structural phase transitions, then the interface must represent a barrier that hinders heat and mass transport between the upper and lower mantle. Thus the dynamic state and the thermal structure of the Earth's mantle can be rather different depending upon the composition of the lower mantle. The electrical conductivity of candidate mantle mineral assemblages can be an important discriminator between the various models of lower-mantle compositions, as there are independent estimates of the electrical conductivity of the mantle.

Induced currents

Observations of the geomagnetic field constrain the electrical conductivity of the lower mantle by two distinct methods⁴. First, external magnetic fields induce currents in the Earth, which in turn produce secondary magnetic fields. From the analysis of the variation of these secondary magnetic fields through time, bounds can be put upon the electrical conductivity of the Earth's mantle. Second, the attenuation of high-frequency components of the internally generated magnetic field provides information on the electrical conductivity of the lower portions of the mantle. Of course the latter technique requires some assumptions about the original frequency content of the internal field.

Previous studies⁵⁻⁷ of the electrical conductivity of candidate mantle mineral assemblages have led to differing conclusions about the composition of the lower

mantle. Li and Jeanloz^{5,6} concluded that the lower mantle is either not as conducting as geomagnetic results indicate or that something other than magnesium-iron silicate perovskite or perovskite-dominated assemblages of upper-mantle composition determines the electrical conductivity of the lower mantle. Peyronneau and Poirier⁷ measured the conductivity of both perovskite and perovskite-dominated assemblages and determined that the conductivities of their samples match the geomagnetic constraints. Thus the lower mantle could have the same composition as the upper mantle.

Li and Jeanloz have since measured the electrical conductivity of several magnesio-wüstites with differing iron contents⁸ and of perovskite and perovskite-dominated assemblages with a higher iron content than their original samples¹. Their results of magnesio-wüstite are in very good agreement with the data of Mao and Bell⁹. This is a significant cross check, as Mao and Bell used techniques similar to Peyronneau and Poirier.

The newest results¹ indicate that magnesio-wüstite's electrical conductivity is a strong function of iron content and that the electrical conductivity of the more iron-rich perovskite and perovskite assemblages is compatible with the electrical conductivities that are obtained from geomagnetic studies. Li and Jeanloz conclude from this, in opposition to Peyronneau and Poirier, that the Earth's lower mantle is enriched in iron relative to the upper mantle. The latest addition to this growing body of literature is the study by Wood and Nell², who have measured the electrical conductivity of magnesio-wüstite as a function of temperature and oxygen fugacity (roughly, partial pressure) at atmospheric pressure, and find in favour of Peyronneau and Poirier.

To see why the story seems to change so much, a brief discussion of the experimental techniques is in order. Both Li and Jeanloz, and Peyronneau and Poirier use a laser-heated diamond anvil to transform their starting material to the lower-mantle assemblages of interest. Li and Jeanloz use the same laser to heat their samples when measuring the electrical conductivity at high pressures. Peyronneau and Poirier, and Mao and Bell, on the other hand, use an external furnace that is placed around the diamond anvils to heat the sample and gasket assembly; they enjoy smaller temperature gradients and have greater control of the temperature of their samples than Li and Jeanloz, but also have to live with a smaller range of temperatures and pressures (298–700 K, and 52 gigapascals, compared with 4,500 K and 82 gigapascals).

An external heater limits the pressure range because the diamond anvils and their support apparatus are heated along with the

sample: with laser heating, only the sample is heated. The upper limit of temperature for the externally heated diamond anvil cell is controlled by the oxidation of the outer surfaces of the diamond anvils (which are at ambient pressure). Thus Peyronneau and Poirier must extrapolate their results in both pressure and temperature to compare them to the geomagnetic results.

Wood and Nell's measurements on magnesio-wüstite were made at atmospheric pressure and high temperature under controlled conditions of oxygen fugacity. It should be noted that the oxygen fugacity in a laser-heated diamond-anvil-cell sample is buffered by cold sample material, unless of course the diamond anvils are damaged and strongly reducing carbon enters into the sample. (All the samples are representative of anhydrous conditions. Li and Jeanloz³ have recently demonstrated that volatiles can increase the electrical conductivity of samples at lower mantle conditions.)

Activation energy

Although Li and Jeanloz disagree with Peyronneau and Poirier over the absolute values of the electrical conductivity, they agree that ionic conductivity is not an important mechanism for electrical conductivity in the Earth's mantle. The activation energy for conduction that each group measures is about a few tenths of an electronvolt. Furthermore, both find that increasing the pressure increases the conductivity. Each observation supports a mechanism of conduction by charge hopping between Fe²⁺ and Fe³⁺ ions in the solids, and rules out ionic conductivity, which would have a larger activation energy and would lead to decreased conductivity at high pressures.

Wood and Nell show that the electrical conductivity of magnesio-wüstite depends strongly on the oxidation state of the samples, which would affect the proportions of Fe²⁺ and Fe³⁺. Notably, they add experimental support to Li and Jeanloz's supposition¹ that the oxygen fugacity of the samples could explain the difference between the various studies. Paradoxically, Wood and Nell agree with Li and Jeanloz⁸ that magnesio-wüstite dominates the electrical conductivity of the lower mantle, but they also claim to agree with Peyronneau and Poirier that the lower mantle has a similar composition to the upper mantle. But their extrapolation uses the high-pressure values of Peyronneau and Poirier, and Li and Jeanloz⁸ for the pressure dependence of the electrical conductivity: Mao and Bell⁹ show that the electrical con-

1. Li, X. & Jeanloz, R. *J. geophys. Res.* **96**, 6113–6120 (1991).
2. Wood, B. J. & Nell, L. *Nature* **351**, 309–311 (1991).
3. Li, X. & Jeanloz, R. *Nature* **350**, 332–334 (1991).
4. Parkinson, W. D. *Surv. Geophys.* **9**, 235–243 (1988).
5. Li, X. & Jeanloz, R. *Geophys. Res. Lett.* **14**, 1075–1078 (1987).
6. Li, X. & Jeanloz, R. *J. geophys. Res.* **95**, 5067–5078 (1990).
7. Peyronneau, J. & Poirier, J. *Nature* **342**, 537–539 (1989).
8. Li, X. & Jeanloz, R. *J. geophys. Res.* **95**, 21609–21612 (1990).
9. Mao, H. K. & Bell, P. M. in *High Pressure Research: Applications in Geophysics* (eds Manghni, M. H. & Akimoto, S.) 493–502 (Academic, New York, 1977).

ductivity of magnesiowüstites varies rapidly with pressure at low pressures and that the pressure derivative approaches a constant value at higher pressures, in agreement with Li and Jeanloz and Peyronneau and Poirier.

Thus Wood and Nell really do not agree with Peyronneau and Poirier, but would actually find an even higher electrical conductivity for the lower mantle if they used the pressure variations observed by Mao and Bell⁹. The possible solution to this dilemma is that, by measuring the electrical conductivity

at high temperatures (1,173–1,773 K), they have activated ionic conductivity (X. Li, personal communication) which would cause them to over-estimate the electrical conductivity at lower-mantle pressures. Further experiments upon their samples, both at room temperature and as a function of pressure, should help clarify this issue. □

Dion L. Heinz is in the Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA.

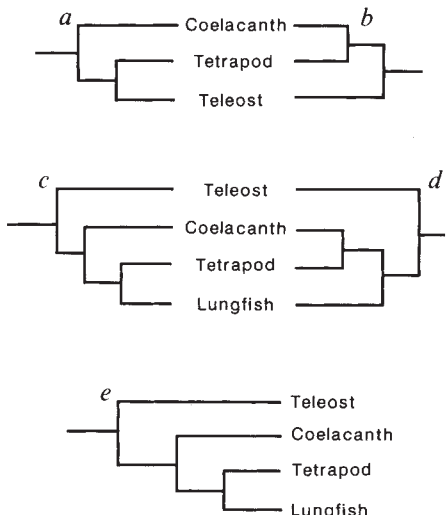
EVOLUTION

Blood lines of the coelacanth

Peter Forey

ON page 394 of this issue, Gorr *et al.*¹ describe the primary structure of the haemoglobin of the coelacanth (*Latimeria*) and suggest that evidence contained in the sequence of amino acids favours the idea that the coelacanth is the closest living relative of tetrapods. This result supports a traditional view, based on palaeontological work², of the relationships of the bony fishes; but it runs counter to other, recently described anatomical³ and molecular data⁴. A simplified classification, including bony fishes, is given in the table.

For the past 25 years amino-acid products of globin genes have been used to reconstruct



The top four trees (a–d) represent the conclusions of cladistic analyses of the relationships between the coelacanth, tetrapods, bony fishes and lungfish using four data sets. a, 115 amino acids of α and β parvalbumin¹⁰ (lungfish excluded). b, 299 nucleotides within the 28S ribosomal DNA¹¹ (lungfish excluded). c, 85 variable nucleotides in mitochondrial DNA genes coding for 12S ribosomal RNA and cytochrome *b* (ref. 4). d, 290 amino acids of α - and β -haemoglobin¹. The bottom tree (e) is a PAUP computer program analysis of mitochondrial DNA and haemoglobin rooted on teleost. The length of this tree=542, consistency index=0.934. The teleost was a 'hybrid', using nucleic acid sequences from *Cichlasoma citrinellum*⁴ and haemoglobin from *Thunnus thynnus*⁴

phylogenetic pathways⁵ and to draw conclusions about rates of molecular evolution. The products of the haemoglobin genes have been widely used and the principal features of their history within vertebrates are now known. Before the appearance of the first gnathostomes (jawed vertebrates) the ancestral haemoglobin gene split by gene duplication and subsequent divergence to yield two gene lineages resulting in the production of an α -haemoglobin chain and a β -haemoglobin chain. Further gene duplications yielded embryonic and adult α - and β -haemoglobins, which operate at different stages of life history, and within mammals there is evidence of further gene duplications. Tracing the history of gene duplication is not always easy, however, because it involves drawing inferences from molecular phylogenetic trees and inferences about homology between haemoglobins in different animals.

Gorr *et al.* demonstrate that the α -chains (143 amino acids) and β -chains (147 amino acids) of all the bony fishes they examined — four teleosts, one lungfish and the coelacanth — show a greater overall similarity with tadpole haemoglobins than with adult amphibian haemoglobins. Nothing is known directly about globin genes in fishes, so it is not certain whether the fish haemoglobins match the larval or the adult amphibian haemoglobins, or neither. The implication here is that the fish haemoglobin matches that of tadpoles, which supports previous studies using cladistic methods of sequence analysis⁶. One explanation for this match is that the gene duplications leading to larval and adult haemoglobins took place in the stem tetrapod, with the adult α - and β -chains diverging more rapidly than the larval chains. Alternatively, the larval and adult gene duplications may have preceded the origin of bony fishes, with the expression of the adult haemoglobins being suppressed in fishes. This second view may be supported by theoretical predictions about the age of divergence of larval and adult chains as measured by the degree of structural difference between them. Gorr *et al.* favour this interpretation and propose that bony fishes

SIMPLIFIED CLASSIFICATION OF VERTEBRATES, EXCLUDING EXTINCT GROUPS

AGNATHA: jawless vertebrates; lampreys and hagfish
GNATHOSTOMATA: vertebrates with jaws
CHONDRICHTHYES: cartilaginous fishes; sharks, rays etc.
OSTEICHTHYES: bony fishes
ACTINOPTERYGII: ray-finned bony fishes
Chondrostei: bichir (*Polypterus*), sturgeon (*Acipenser*) and allies
Neopterygii: modern ray-finned fishes
Ginglymodi: garpike (*Lepisosteus*)
Halecomorphi: bowfin (*Amia*)
Teleostei: majority of extant fishes (e.g. herring, cod, perch)
SARCOPTERYGII: lobe-finned bony fishes
Actinistia: coelacanth (*Latimeria*)
Dipnoi: lungfishes
Tetrapoda: land vertebrates
Lissamphibia: modern amphibians
Anura: frogs and toads
Urodela: newts and salamanders
Amniota: all other land vertebrates

express the larval form of haemoglobin as adults.

The authors also found, by phenetic methods of analysis, that the β -chain shows *Latimeria* may be most closely related to tadpoles whereas the α -chain favours teleosts as immediate tadpole relatives. Their explanation for this apparent anomaly is that the α -haemoglobin molecule has evolved at different rates — slowly in teleosts, rapidly in lungfishes and coelacanths. This darwinian explanation of irregularity is supported by some⁵, but not by others⁷ who believe that the same molecule evolves at the same rate across species lineages. One way in which this idea may be further explored is to examine the haemoglobin of primitive ray-finned bony fishes such as the bichir (*Polypterus*), the sturgeon (*Acipenser*) and the garpike (*Lepisosteus*) (see table).

Cladistic analyses give Gorr *et al.* the same signal for α - and β -chains and places *Latimeria* as the sister-group of tadpoles. Gene-lineage trees such as these may differ from one another, and from species-lineage trees based on anatomical data, for a number of reasons — for example the confusing effects of gene duplication, gene conversion, possible unequal rates of molecular evolution and back amino-acid substitution, as well as difficulties for the investigators when

- Gorr, T., Kleinschmidt, T. & Fricke, H. *Nature* **351**, 394–397 (1991).
- Schultze, H.P. *J. Morph. Suppl.* **1**, 39–74 (1987).
- Rosen, D.E., Forey, P.L., Gardiner, B.G. & Patterson, C. *Bull. Am. Mus. nat. Hist.* **167**, 159–276 (1981).
- Meyer, A. & Wilson, A.C. *J. molec. Evol.* **31**, 359–364 (1990).
- Goodman, M., Romero-Herrera, A.E., Dene, H., Czelusniak, J. & Tashian, E. in *Macromolecular Sequences in Systematic and Evolutionary Biology* (ed. Goodman, M.) 115–191 (Plenum, New York, 1982).
- Goodman, M., Miyamoto, M.M. & Czelusniak, J. in *Molecules and Morphology in Evolution: Conflict or Compromise?* (ed. Patterson, C.) 141–176 (Cambridge Univ. Press, 1987).
- Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
- Czelusniak, J., Goodman, M., Moncrief, N.D. & Kehoe, S.M. *Meth. Enzymol.* **183**, 601–615 (1990).
- Kluge, A. *Syst. Zool.* **38**, 7–25 (1989).
- Maeda, N., Zhu, D. & Fitch, W. *Molec. Biol. Evol.* **1**, 473–488 (1984).
- Hillis, D.M. & Dixon, M.T. in *The Hierarchy of Life* (eds Fernholm, B., Bremer, K. & Jörmwall, H.) 355–367 (Excerpta Medica, Amsterdam, 1989).

aligning sequences before their analysis.

A technique for collating disparate information is to determine the degree to which phylogenetic trees based on different molecular data agree with each other. But when it comes to the relationships of bony fishes, there are very few sets of data which embrace the same key taxa; usually, information for one or another is missing. The figure gives a summary of the published results of separate cladistic analyses using various molecular data. The common element in these trees simply results in a three-way split between coelacanth, lungfish and tetrapod; that is, neither lungfish nor coelacanth can be judged closer to tetrapods. Another technique is to put together all molecular data in the same data matrix^{8,9} (tandemly align various molecular sequences). This is done here using a total of 375 potential characters (290 amino acids — 57 amino-acid sites are taxo-

nomically informative — used by Gorr *et al.*, and 85 nucleotides for mitochondrial DNA coding for 12S ribosomal RNA and cytochrome *b*) to yield *e* in the figure. The result shows the lungfishes as the nearest relatives of tetrapods — the nearest competing tree, which places *Latimeria* and lungfishes as sister-groups, requires seven extra character transformations.

This does not necessarily make the 'combined data' solution correct; rather it goes some way towards summarizing the current molecular data. Gorr *et al.*'s results, correct for haemoglobin, will surely prompt the search for more sequences and, one hopes, investigation of the nature of fish globin genes. □

Peter Forey is in the Department of Palaeontology, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

NUCLEOSYNTHESIS

Finding the proton drip line

Grant J. Mathews

Six new nuclear isotopes near the limits of stability for proton-rich nuclei in the region of gallium through to strontium have been discovered¹ by a group working at Michigan State University. These isotopes may redefine the region of known stable proton-rich nuclei known as the 'proton drip line'. They also may imply a different evolution for models describing the generation of astronomical X-ray bursts from high-temperature thermonuclear hydrogen burning on accreting neutron stars²⁻⁵.

The interest in these new isotopes stems partly from the fact that their very existence serves to map out the delicate balance between the attractive nuclear force and repulsive electromagnetic force in atomic nuclei. The nuclear force is attractive between a proton and a neutron but slightly repulsive between two protons or two neutrons. Thus, there is a limit to how many more protons than neutrons (or vice versa)

one can add to a nucleus. This situation is further aggravated by the electromagnetic Coulomb repulsion among protons which strives to push the nucleus apart. The limits to how many protons or neutrons can be added are known as the proton and neutron drip lines. The newly made isotopes have nearly equal numbers of protons and neutrons (*Z* and *N*, respectively) and so have the strongest nuclear binding. However, they have so many protons that the Coulomb repulsion nearly cancels the nuclear force. Thus, these nuclei are delicately bound and it is difficult to predict theoretically which nuclei will be bound and which not.

One way to make such isotopes is to collide together a heavier proton-rich nucleus and a lighter nucleus. As the two nuclei interact, nucleons will be stripped away from the heavy nucleus producing new isotopes which will recoil from the momentum of the collision. In the new experiment, the proton-rich

isotopes were made by bombarding a ⁵⁸Ni target with high-energy ⁷⁸Kr ions at the US National Superconducting Cyclotron Laboratory. The recoiling products were collected and focused through a sequence of magnetic lenses and their masses analysed.

The resulting mass spectra show clear evidence of six previously unknown isotopes: ⁶¹Ga, ⁶²Ge, ⁶³Ge, ⁶⁵As, ⁶⁹Br and ⁷⁵Sr, as well as other known isotopes (see figure). Two of these nuclides, ⁶⁵As and ⁶⁹Br are of particular interest for two reasons. Previous theoretical mass extrapolations⁶ could not show whether these nuclides do or do not have bound states for the last proton. The new observation thus extends the known⁷ proton drip line and places severe constraints on models for the properties of symmetric (*N* ≈ *Z*) nuclei at the limits of stability.

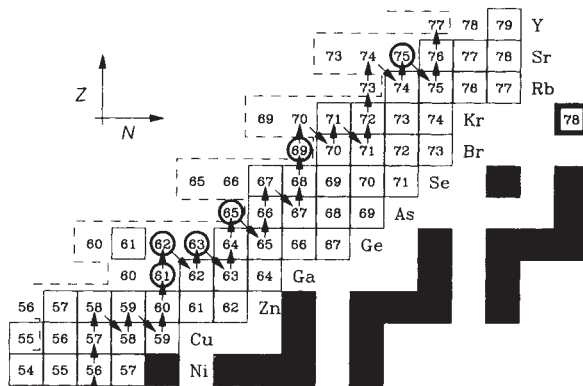
Perhaps of even more interest is the implication of these new data for the thermonuclear burning of hydrogen at temperatures in excess of a billion (10⁹) kelvin (refs 2-5). High temperatures occur naturally in the cores of massive stars at the end of their lifetime but these are very depleted in hydrogen. A high-temperature hydrogen-rich environment requires a more exotic setting such as the thermonuclear bursts associated with novae, supernova shocks, accreting white dwarfs and neutron stars — and even the Big Bang according to models⁸ in which primordial nuclear matter is inhomogeneously distributed.

The highest temperatures are described in models which attempt to explain astronomical type I X-ray bursts⁵ in terms of material which has gradually accreted onto the surface of a neutron star from a nearby companion star: once the pressure and density of the accreted material exceed the threshold for nuclear reactions, the thermonuclear energy quickly heats the material to high temperature in the intense gravitational field of the neutron star. The thermonuclear burning in this environment is then characterized by rapid proton captures (called the rp-process²) from C, N and O nuclei up to the proton drip line.

Of particular interest in the context of the new measurements is that the rp-process would cease⁹ when the nucleus ⁶⁴Ge was formed if ⁶⁵As could not be made by adding another proton to ⁶⁴Ge. The half life for the β⁺-decay of ⁶⁴Ge is long (64 seconds) compared to the timescale for the thermonuclear burst. Therefore, the rp-process would have time to cool below the threshold for hydrogen burning as it waited for the decay of ⁶⁴Ge. The fact that ⁶⁵As and ⁶⁹Br are now known to be stable against the loss of a single nucleon may allow the rp-process to proceed beyond ⁶⁴Ge, through ⁶⁹Br and on to heavier nuclei (see figure). As ⁷³Rb was not observed in the reported measurements it is probably beyond the proton drip line. The new waiting point for the rp-process is, therefore, probably the 17-second decay half life of ⁷²Kr.

All of this means that the energy could be generated in the rp-process at a faster rate,

Nuclear chart for the region studied by Mohar *et al.*¹, and the rp-process. The nuclear charge *Z* (the number of protons) increases up the chart; the number of neutrons *N* increases from left to right. The filled squares represent stable nuclei, and the open squares, previously known radioactive isotopes whose mass number (equals *N*+*Z*) is given. The newly discovered isotopes are indicated by the encircled numbers. Astrophysical rp-nucleosynthesis occurs by the absorption of a proton (vertical arrows indicating the increase of *Z* by 1) followed by radioactive β⁺-decay (emission of a positron, diagonal arrows). If ⁶⁵As were unstable (to the left of the proton drip line, indicated by the broken line), the rp-process would terminate at ⁶⁴Ge, which would fail to absorb a proton. The new results suggest it could continue as far as ⁷⁴Rb. (From ref. 1.)



and the available hydrogen would be exhausted more completely than in previous models for X-ray bursts^{2,3}. The more complete burning of hydrogen may be a problem when it comes to explaining one of the puzzling features of X-ray bursters. In a few cases, multiple bursts are observed⁵ 5–10 minutes apart, a quiescent period far too short for accretion to build up a critical mass. It has been suggested^{3,10} that these repeated X-ray flashes are due to a re-ignition of hydrogen remaining when the rp-process ceases. Such a re-ignition might be caused by a convective instability. But if the burning of hydrogen during the first burst is more complete, ignition of the successive bursts seems less likely.

Even so, we cannot yet say for certain what effect, if any, these new isotopes will have on the models. The observation of particle-stable ⁶⁵As and ⁶⁹Br in this experiment does not guarantee that these nuclei are stable enough to contribute to the rp-process. Even if ⁶⁵As is bound, the binding energy of its last proton would need to be more than a few hundred kiloelectronvolts for it to be a factor in X-ray bursts. Otherwise, the excitation of ⁶⁵As by background thermal photons would photodissociate the nucleus before it decayed to ⁶⁵Ge and allow the rp-process to continue. The same is true for ⁶⁹Br. But the fact that very few ⁶⁵As and ⁶⁹Br events were observed suggests that these nuclei may be only very weakly bound.

In the end, we will only be able to quantify the influence of these new isotopes on models for X-ray bursts by accurately measuring their binding energies. This could be done either from the β^+ -decay endpoint energy (something which would be difficult indeed), or by producing these nuclei in two-body reactions using radioactive ion beams¹¹ (for example, the reaction $^{58}\text{Ni} + ^{10}\text{C} \rightarrow ^3\text{H} + ^{65}\text{As}$) and measuring the product recoil energy. Such measurements, however, will probably have to wait until a new generation of radioactive ion-beam accelerators is built. □

Grant J. Mathews is at the Lawrence Livermore National Laboratory, University of California, Livermore, California 94550, USA.

The semiotics of charge

N. Michael Green

ON page 414 of this issue¹, a group from Richard Klausner's laboratory describe results that will further constrain speculation about the function of charged residues in transmembrane segments. The paper, by Cosson *et al.*, is the latest of several that substantiate the importance of such residues in

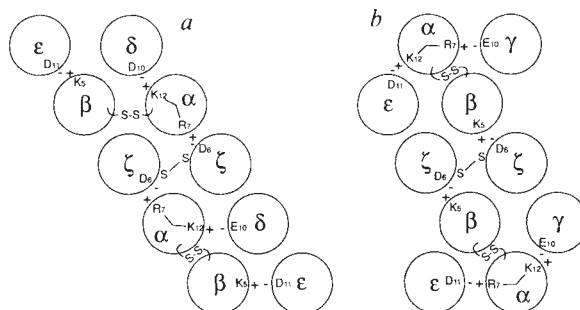
Cosson and colleagues now provide further evidence that the charged residues have a role in stabilizing the complex. The authors made a number of artificial constructs in which the transmembrane segment of TCR α was fused to an extramembranous marker (the α chain of the interleukin-2 receptor,

IL2R, also known as Tac), and its positive charges were replaced by hydrophobic residues. The constructs were co-expressed with the negatively charged CD3- δ and complex formation was assayed immunologically. Stable complexes were formed provided that either one of the two positive charges at positions five and ten remained, thereby confirming results using less direct assay methods⁸. In further experiments, positive charges were introduced at different sites in the native, uncharged helix of IL2R. These formed complexes with CD3- δ provided that the arginine was approximately eight residues inside the hydrophobic segment, matching the position of the negative charge of CD3- δ . The implication is that charge interactions are sufficient to stabilize the complex, because the sequence of the hydrophobic residues in IL2R is quite different from that in TCR α .

Earlier experiments from

Klausner's laboratory, using a similar system, had shown that as well as stabilizing the complex the charges act as signals which determine the fate of any TCR α subunit produced in excess of the limiting amount of the ζ chain⁹. Similar results have been obtained with TCR β ¹⁰. The excess of TCR α or β subunits or CD3- δ subunits was degraded in a pre-Golgi pathway, probably in a separate compartment from the endoplasmic reticulum (ER) itself, although often referred to as 'ER'. It seemed that in the absence of the appropriate subunits with complementary charges, TCR α or β were removed for degradation. The charge was an essential signal for the pre-Golgi pathway. Without it the protein appeared on the plasma membrane (the default pathway), whereas truncation of the whole hydrophobic segment caused retention in the ER without degradation. Corresponding experiments on the role of the negative charge in determining the fate of CD3- δ have not been reported.

Given the dominant role of transmembrane charge in determining the interactions of these membrane-bound subunits, what



Possible arrangements of transmembrane helices of the TCR-CD3 complex which optimize charge interactions. The main features of these proposals are that the disulphide-linked ζ chains are symmetrically situated with respect to either α (a in the figure) or β (b), linking two ζ units in a dimer; and the allocation of γ and δ to separate complexes¹¹ (the allocations shown are arbitrary and could be reversed). In this way it is possible to maintain electroneutrality and to position complementary charges at approximately matching depths within the membrane, as indicated by the numbers associated with Lys (K), Arg (R), Asp (D) and Glu (E). In a the only serious mismatch is between K5 (β) and D11 (ϵ), which could just be bridged by extending the lysine side chain or by inserting the β subunit farther into the membrane. In b this mismatch is avoided, but the model is inconsistent with the evidence for interaction between α and δ (ref. 12). The depths assigned to the charged residues differ slightly from those used in ref. 1, emphasizing the arbitrary element in these assignments. The disulphide-linked cysteines of the ζ chain are four residues N-terminal to the aspartic acid residues (D6), which are therefore offset slightly (40°).

stabilizing the interactions between proteins.

An early model of bacterial rhodopsin proposed that the charges might form stabilizing ion pairs². But the structure itself³ shows that the positive charges are probably neutralized in the polar head group region of the lipids, whereas the negative charges are accommodated either by protonation or by location in a region accessible to solvent. Clearly, proteins have many ways of accommodating buried charges⁴.

In proteins anchored by a single transmembrane segment, options are limited and intramembranous charges are uncommon⁵; in fact there were no examples before arginine and lysine were found in the α and β chains of the T-cell receptor (TCR α and TCR β)⁶. Because the receptor is oligomeric, it was suggested that the charges were buried between subunits. This interpretation became even more plausible when the sequences of the γ and δ chains of the CD3 complex, which provides the signalling unit for the clone-specific $\alpha\beta$ dimer, showed a single negative charge in each hydrophobic segment⁷.

- Mohar, M. F. *et al.* *Phys. Rev. Lett.* **66**, 1571–1574 (1991).
- Wallace, R. K. & Woosley, S. E. *Astrophys. J. Suppl.* **45**, 389–420 (1981).
- Ayasli, S. & Joss, P. C. *Astrophys. J.* **256**, 637–651 (1982).
- Hanawa, T., Sugimoto, D. & Hashimoto, M. *Publ. Astr. Soc. Japan* **35**, 491–506 (1983).
- Lewin, W. H. G. & Joss, P. C. in *Accretion Driven Stellar X-Ray Sources* (eds Lewin, W. & van den Heuvel, E.) 41–115 (Cambridge Univ. Press, 1983).
- Hauastine, P. E. (ed.) *1986–1987 Atomic Mass Predictions (Atom. Data nucl. Data Tables* **39**, 185–393; 1988).
- Pougheon, F. *et al.* *Z. Phys. A* **327**, 17–24 (1987).
- Kajino, T., Mathews, G. J. & Fuller, G. M. *Astrophys. J.* **364**, 7–14 (1990).
- Wallace, R. K. & Woosley, S. E. in *High Energy Transients in Astrophysics* (ed. Woosley, S. E.) 319–324 (Am. Inst. Phys., New York, 1984).
- Woosley, S. E. & Weaver, T. A. in *High Energy Transients in Astrophysics* (ed. Woosley, S. E.) 273–297 (Am. Inst. Phys., New York, 1984).
- Mathews, G. J. in *Proc. Workshop Sci. Intense Radioactive Ion Beams* (eds McClelland, J. B. & Vieira, D. J.) 213–230 (Los Alamos National Laboratory, 1990).

might be the arrangements of the trans-membrane helices which maximize charge interactions? A particularly simple solution emerges (see figure) if one takes into account evidence that the sequentially similar CD3- γ and CD3- δ subunits form distinct isomorphous complexes¹¹. If each CD3 unit includes ϵ , ζ and either γ or δ , then a neutral complex can be formed because the two positive charges on TCR α are five residues apart and would point in almost opposite directions. The models shown are an arbitrary selection from many possible configurations. The rapidly accumulating evidence on pairwise interactions between chains expressed in artificial systems (see, for example, ref. 12) has not been fully taken into account and the new evidence may eliminate most or even all of these suggested arrangements. But the diagram shows a number of structural arguments which should be considered in the model-building exercises that will become feasible when the overall stoichiometry is more clearly established.

Other systems in which charged trans-membrane anchors have an important role include the IgE receptor¹³ and membrane fusion induced by the human immunodeficiency virus¹⁴. The experimental approach of combining site mutagenesis with expression of chimaeric constructs is a powerful method for analysing these systems, which are difficult to obtain in sufficient quantity for conventional studies.

Nevertheless, caution is required in interpreting the significance of complex formation in detergent, because in principle detergent could allow interchange between subunits from different membranes. Experimental studies to define the effect of charges in simpler systems are also desirable — for example, such studies could amplify the observations that although a foreign charge in a viral membrane anchor may not prevent insertion, it does lead to elimination through the lysosomes rather than assembly^{15,16}. □

N. Michael Green is in the Laboratory of Protein Structure, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

Do little things matter?

Geoffrey King and Armando Cisternas

SINCE geologists discovered both faults and folds in rocks, the relations between the two have been controversial. The focus of this debate, which started in the nineteenth century, is whether faults are a secondary consequence of ductile deformation or ductile deformation a secondary consequence of faulting. As shown by a study by J. Walsh, J. Watterson and G. Yielding on the extensional history of the North Sea, elsewhere in this issue (*Nature* **351**, 391–393; 1991), and an earlier study by C. H. Scholz and P. A. Cowie (*Nature* **346**, 837–839; 1990), the debate is very much alive.

A fundamental problem in geology concerns our ability to build models of structural change. Fashion has sometimes favoured models in which ductile behaviour predominates, at other times the vogue has been for those in which rigid blocks move between major faults. Because it has proved difficult, in practice, to produce realistic models that incorporate both features, imaginative effort has been expended in attempting to determine which of the two best approximates the behaviour of the Earth at various scales and various depth ranges.

Among geophysicists, though not most structural geologists, plate tectonics, by demonstrating the occurrence of block motion at the largest scale, dramatically shifted preconceptions towards concepts of blocks. The political analogy of 'buffer plates' that deform in a quasi-continuous fashion in response to the motions of larger and more powerful blocks started the idea that, for continental deformation, continuum models might prove to be better than rigid plates in some places. The dramatic laboratory model of a rigid indenter, representing the northward-moving Indian continent, extensively deforming southern Asia, made this new paradigm respectable. However, it has remained unclear whether this deformation is truly ductile, or whether it represents motion distributed over numerous faults.

Fractal descriptions entered this debate about 10 years ago, when it became clear that some of the finest examples of natural fractals exist in geological and geomorphological features and in the behaviour of earthquakes. Faults can look much the same over scales from centimetres to tens or even hundreds of kilometres, and the range of scales over which earthquakes behave in a broadly self-similar fashion exceeds six decades. Could ductile deformation be the consequence of self-similar fault motion below any chosen scale of observation? The answer emerging is "sometimes and at some scales".

In California, if earthquakes are placed in fault-size classes with characteristic fault lengths varying by factors of two, the sum of all events smaller than any given class gives

the same deformation as that class itself. In other words, smaller events are important. This happy result breaks down for the biggest events, those that contribute the most deformation. They are simply too big. Just where a fractal relation would be most useful, it does not work. The same appears to be true outside California and has led to the general statement that little earthquakes can be ignored. With this view in mind, the rate at which big earthquakes occur has been compared with that predicted by plate rates and declared to be too small. The implication that much motion does not occur on earthquake faults seems irresistible. Perhaps the missing deformation occurs on creeping faults as continuum deformation.

In principle, the scaling relations of degree of slip and dimensions of geological faults should resolve this issue. Geological faults include faults that move in earthquakes and those that slip by creep. But researchers determining the partitioning of geological motion between large and small faults appear to be stumbling on problems analogous to those for seismic faults. Scholz and Cowie have found data suggesting that the contribution of smaller faults can be neglected, but Walsh *et al.* find the opposite. As popular models for the lithospheric stretching needed to explain the amount the floor of the North Sea has subsided suggest that there should be much greater deformation than can be provided by the largest faults, the first interpretation supports fractal block motion and the latter a large role for continuum deformation.

It may be too early to take sides in this debate, particularly for those of us who lack a strong partiality in favour of one conclusion rather than the other. Although the data sets employed are greatly superior to those available only a decade ago, they are very far from complete. Nor can we be sure that our stretching models for the evolution of sedimentary basins provide a sure measure of extension. In the case of comparing seismicity with plate rates we must realize that seismicity rates have varied enormously within the time of instrumental recording and in historical time. We have no reason to suppose that such irregularity does not exist over even longer time spans. Thus comparisons with plate rates determined from magnetic reversals could be spurious. In resolving the question of continuum versus block deformation our tools are really very blunt. Consequently, it seems likely that this question, so fundamental to understanding the physics of the ground beneath our feet, will still be with us in the next century. □

Geoffrey King and Armando Cisternas are at the Institut de Physique du Globe, 5 rue René Descartes, 67084 Strasbourg, France.

1. Cosson, P., Lankford, S. P., Bonifacio, J. S. & Klausner, R. D. *Nature* **351**, 414–416 (1991).
2. Engelman, D. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2023–2027 (1980).
3. Henderson, R. *et al. J. molec. Biol.* **213**, 899–929 (1990).
4. Quiocho, F. A., Sack, J. S. & Vyas, N. K. *Nature* **329**, 561–563 (1987).
5. van Heijne, G. *Curr. Top. Membr. Trans.* **24**, 151–179 (1985).
6. Hedrick, S. M. *et al. Nature* **308**, 153–158 (1984).
7. van den Elsen, P. *et al. Nature* **314**, 542–546 (1985).
8. Blumberg, R. S. *et al. J. biol. Chem.* **265**, 14036–14043 (1990).
9. Bonifacio, J. S., Cosson, P. & Klausner, R. D. *Cell* **63**, 503–513 (1990).
10. Willeman, T. *et al. Cell Reg.* **1**, 907–909 (1990).
11. Alarcon, B. *et al. EMBO J.* **10**, 903–912 (1991).
12. Manolios, N., Bonifacio, J. S. & Klausner, R. D. *Science* **249**, 274–276 (1990).
13. Blank, U. *et al. Nature* **337**, 187–189 (1989).
14. Helseth, E. *et al. J. Virol.* **64**, 6314–6318 (1990).
15. Adams, G. A. & Rose, J. K. *Molec. Cell Biol.* **5**, 1442–1448 (1985).
16. Davis, G. L. & Hunter, E. J. *Cell Biol.* **105**, 1191–1203 (1987).

The ring on a finger

Simon van Heyningen

THE overwhelming diarrhoea of cholera is produced by a protein toxin secreted by *Vibrio cholerae* bound to the brush border of the intestine. Some strains of *Escherichia coli* produce a similar, though less serious disease, and a so-called 'heat-labile toxin' that is essentially the same protein: they are about 80 per cent similar in their amino-acid sequences and have the same mechanism of action. After years of remarkably successful speculation, we now know the structure of the heat-labile toxin from an X-ray diffraction study reported on page 371 of this issue¹ by Sixma *et al.* from Groningen. The toxin is made up of two quite different subunits that can fold and function independently, yet are tightly associated in the intact protein: five copies of the cell-binding B subunit make a ring upon a finger sticking out from the single enzymic A subunit (shown in Fig. 3 of the paper on page 375). No doubt the structure of cholera toxin is the same in all but detail.

Although most of the published work on these proteins was done on cholera toxin, and a preliminary diffraction paper was published 14 years ago², there was little further progress until the Dutch group decided to concentrate on heat-labile toxin. They made heroic amounts of it, and grew an astonishing array of crystals³. Even so, some early crystals were irritatingly non-isomorphous; the group's first useful data sets were collected from the same crystal before and after soaking with heavy metal.

Subunits

The A and B subunits of these toxins (like similar components of many others) have different tasks. The five B subunits, each of relative molecular mass about 11,600 (M_r 11.6K), bind to ganglioside GM1, which is found in the outer membrane of essentially all eukaryotic cells (the toxins have very little cell specificity). Binding leads to the proteolytic 'nicking' of subunit A into two polypeptide chains, A1 (M_r 22K) and A2 (M_r 5K), and entry of the A1 chain into the cell. The A1 polypeptide catalyses the ADP-ribosylation of the G protein G_{α} , leading to activation of adenylate cyclase, and, presumably, ultimately to the diarrhoea, although the connection has never been established experimentally.

The subunit structure of cholera toxin was determined nearly 20 years ago, and it was clear then from symmetry and other considerations that its five B subunits had to be arranged in a ring with the A subunit sitting on top of them. Many authors produced

diagrams showing such a structure⁴, and more recent support came from two-dimensional diffraction experiments⁵. It is gratifying for traditional protein chemists to see that the old arguments were right, but no one could have predicted the apparently unique interlocking secondary structure.

The ring of B subunits is held together in a way that the authors rightly call "quite spectacular": an individual subunit has two sets



Vibrio cholerae — its toxin is closely akin to heat-labile toxin.

of β sheet each with three anti-parallel strands, but the first set of one subunit forms a six-stranded sheet with the second set of its neighbour. Each subunit also has a 19-residue α helix, arranged so that the five of them line a pore in the middle of the ring, 30 Å deep and with diameter tapering from 15 Å to 11 Å.

Into this pore fits a hairpin loop comprising the C-terminal half of the A2 peptide (which is known from cross-linking experiments to be involved with binding to the B oligomer). The loop is linked to A1 by the N-terminal half of A2, a 23-residue helix which holds the globular A1 peptide and the B pentamer together but slightly apart like an umbilical cord (see Fig. 3 of the paper). The region of A2 held in the pore is strongly bound to the B pentamer by at least 15 salt bridges. Because five identical subunits bind to a single different subunit, it is hardly surprising that this arrangement is not quite symmetrical. There are many examples of proteins in which association with an asymmetrical partner induces asymmetry in a previously symmetrical molecule.

The mechanism by which A1 is taken into a cell is not clear; it need not be very efficient because cells can bind millions of toxin molecules, but only a few of A1 are enough to ADP-ribosylate all the G protein. An important question is which way up the toxin binds to a cell — is it with subunit A held close

to the membrane, or with the pentamer closest to the membrane and subunit A on the other side? Most of the evidence⁵ points to the first possibility, in which case the surface of the pentamer that would be on the membrane side looks very suitable because it is flat and highly charged. But Sixma *et al.* show that the ganglioside-binding sites are not on this flat, charged surface near subunit A, but on the other side. Perhaps the toxin does actually bind with subunit A away from the membrane, but that would make it difficult for A to get to the cell surface. Alternatively, if the toxin binds with subunit A downwards, a ganglioside molecule might just be long enough to loop all the way round the outside of the ring while remaining anchored in the membrane. Neither possibility is immediately attractive.

Enzyme

The structure of the A1 peptide would have been worth solving in its own right. It is an enzyme that catalyses a post-translation modification used by many other toxins such as diphtheria and pertussis, as well as by ADP-ribosyltransferases whose function is unknown but which are found in most cells⁶. There are some marked similarities in the active sites of heat-labile toxin and of exoenzyme A of *Pseudomonas aeruginosa*⁷, the only other such protein whose structure has been determined;

perhaps they are examples of a new type of NAD^+ -binding domain characteristic of the ADP-ribosyltransferases. Neither toxin has the well-known 'Rossmann' $\beta\alpha\beta\beta$ fold characteristic of the NAD^+ -binding site of dehydrogenases, but this is not surprising because the reaction catalysed by the toxins is quite different.

What good will all this do in South America, Bangladesh and wherever else there is cholera? Sadly, not very much. Protein chemistry might help with the design of better vaccines, but those most at risk from cholera are the poor, to whom even the simple, effective and cheap rehydration therapy⁸ is not always available. What they need is not vaccines, but clean water supplies and better sanitation. □

Simon van Heyningen is in the Department of Biochemistry, University of Edinburgh, George Square, Edinburgh EH8 9XD, UK.

1. Sixma, T. K. *et al.* *Nature* **351**, 371–377 (1991).
2. Sigler, P. S., Drury, M. E., Kiefer, M. C. & Finkelstein, R. A. *Science* **197**, 1277–1279 (1977).
3. Prank, S. E. *et al.* *J. Biol. Chem.* **260**, 13580–13584 (1985).
4. van Heyningen, S. *Biol. Rev.* **52**, 509–549 (1977).
5. Ribb, H. O., Ludwig, D. S., Mercer, K. L., Schoolnik, G. K. & Kornberg, R. D. *Science* **239**, 1272–1324 (1988).
6. Moss, J. & Vaughan, M. (eds) *ADP-ribosylating Toxins and G Proteins: Insights into Signal Transduction* (Am. Soc. Microbiol., Washington, 1990).
7. Brandhuber, B. J., Allured, V. S., Faibel, T. G. & McKay, D. B. *Proteins* **3**, 146–154 (1988).
8. Hirschhorn, N. & Greenough, W. B. *Sci. Am.* **264**, 16–22 (1991).

Etna's greenhouse pump

Terrence Gerlach

MOUNT Etna emits CO₂ gas at an astonishing rate — 25 million tons a year is the conservative estimate of Allard *et al.* on page 387 of this issue¹. This equals the output of four 1,000 MW coal-fired power stations. Active volcanoes on land typically pump out CO₂ at rates more like 0.1–2 million tons a year, with a median rate of approximately 1.3 mil-

lions apparently lost in this way rather than through eruptive degassing. Appreciation of the importance of quiescent CO₂ emissions from volcanoes began to grow several years ago when it became clear^{3,4} that it was the principal mode of gas release from Kilauea in Hawaii. The low solubility of CO₂ in silicate melts at upper-crustal depths, where mag-



Mount Etna — the volcano emits hitherto unsuspected amounts of carbon dioxide, almost half of the non-eruptive degassing taking place through the flanks. (Photo courtesy of P. Allard.)

lion tons a year². Thus, Mount Etna stands out among its subaerial peers. Indeed, its estimated output rivals even that of the entire 60,000-km-long mid-ocean-ridge volcanic system, currently estimated to emit 30–65 million tons of CO₂ annually².

Why is the CO₂ output from Etna so large? We do not know for sure, but one possibility is that Etna's CO₂ output is augmented by the decarbonation of the limestones and dolomites of the platform on which Etna rests. This suggestion, by Allard *et al.*, is based on the presence of carbonate xenoliths in ejecta from Etna and the elevated ¹³C content of CO₂ from more peripheral sites on the volcano, close to values more typical of marine carbonates. But this does not help much in explaining why Etna's SO₂ emission rates are also unusually high. A more likely possibility, in view of abundant petrological evidence for the enrichment of CO₂ in alkaline magma, is that high CO₂ emission rates are characteristic of alkaline volcanic systems. So if Etna is any indication, other alkaline volcanoes (such as Mount Erebus or Nyiragongo) may also vent large CO₂ fluxes.

The results for Etna provide increasing evidence that CO₂ escapes continuously from active volcanoes by quiescent, non-eruptive degassing. In fact most of the CO₂ in magma supplied to active volcanoes is

mas tend to reside before erupting, causes magmas underlying volcanoes to become depleted in CO₂ as it escapes into overlying hydrothermal fluids and atmospheric plumes. The work by Allard and colleagues indicates that almost half of the noneruptive CO₂ loss at Etna is by diffusive transport through the volcano's flanks. Noneruptive CO₂ degassing is probably a widespread phenomenon at active volcanoes and can cause hazards from the build up of CO₂ concentrations in low areas with poor air circulation. The catastrophe at Lake Nyos, Cameroon, in 1986 probably resulted from the noneruptive degassing of CO₂ from alkaline basaltic magma at depth into a crater lake. The lake ultimately overturned producing a lethal CO₂-rich cloud that killed 2,000 people⁵.

There are many reasons why the degassing of large volumes of CO₂ from volcanoes went largely unnoticed for a long time. Being colourless, odourless and relatively noncorrosive, CO₂ has never made quite the same impact on volcano watchers as the puffy steam clouds of H₂O, the nasal, chromatograph-activating sulphurous gases SO₂ and H₂S, or the zipper-, watch- and camera-damaging halogen gases HCl and HF. Furthermore, noneruptive degassing is rarely spectacular. And the high natural background concentrations of CO₂ in the atmos-

phere present a significant obstacle to direct measurement of volcanic CO₂ emission rates by remote spectroscopic techniques similar to the COSPEC technique that is used widely to measure volcanic SO₂ emission rates. Even petrologists who seek the initial, pre-eruptive volatile content of magmas in microscopic bits and pieces of glass formed from melt trapped in crystals have been largely foiled in their search for the initial CO₂: its low solubility in silicate melt and strong preference for taking up residence in a fluid phase, allows it to escape entrapment within the glass inclusions.

The degassing of CO₂ from volcanoes was once an important part of the natural order of things. In pre-industrial times, it played an important role in balancing a potential deficit in the carbon geochemical cycle. Without the resupply of CO₂ by volcanic and metamorphic degassing, removal of atmospheric CO₂ by silicate weathering, carbonate deposition and burial of organic matter would have depleted the CO₂ content of the pre-industrial atmosphere in 10,000 years and the atmosphere-ocean system in 500,000 years^{6,7}. The CO₂ content of the atmosphere-ocean system has varied in the past, but not at the rate expected if CO₂ were removed and not replenished. It is assumed, therefore, that CO₂ degassing from the Earth's interior restored the deficit from surficial processes and balanced the pre-industrial atmospheric CO₂ budget on a timescale of 10⁴–10⁶ years. Earlier atmospheric-balancing calculations^{6,7} imply present-day (pre-industrial) CO₂ degassing rates of 265 million tons a year; more recent calculations⁸ suggest degassing rates may be as high as 485 million tons a year.

The global rate of CO₂ emission from subaerial and submarine volcanoes is uncertain, but a conservative estimate of 130–175 million tons a year has been made². This estimate indicates that volcanoes contribute 35–65 per cent of the CO₂ needed to balance the deficit in the pre-industrial atmosphere-ocean system. If the results of Allard *et al.* are correct, Etna alone supplied a whopping 10 per cent of the missing CO₂ and clearly was a leading player in restoring the deficit in pre-industrial times. People have, of course, now replaced volcanic and metamorphic processes as the restorers of the atmosphere-ocean carbon deficit. Seeing to it that the deficit is made up in the long run is undoubtedly in our best interest, but we may be overdoing it in the short run. Emissions of

1. Allard, P. *et al.* *Nature* **351**, 387–391 (1991).
2. Gerlach, T. M. *Eos* (in the press).
3. Greenland, L. P., Rose, W. I. & Stokes, J. B. *Geochim. cosmochim. Acta* **49**, 125–129 (1985).
4. Gerlach, T. M. & Graeber, E. J. *Nature* **313**, 273–277 (1985).
5. Freeth, S. J. *et al.* *Nature* **348**, 201 (1990).
6. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, New York, 1978).
7. Berner, R. A. *et al.* *Am. J. Sci.* **283**, 641–683 (1983).
8. Berner, R. A. *Geochim. cosmochim. Acta* **54**, 2889–2890 (1990).
9. Boden, T. A. *et al.* *Trends '90 - A Compendium of Data on Global Change* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, 1990).
10. Berner, R. A. & Lasaga, A. C. *Sci. Am.* **74**–81 (March 1989).

CO₂ from fossil-fuel burning, cement production and gas flaring now amount to 22,000 million tons a year⁹. Contributions from our management of the biosphere are less well quantified, but are potentially of the same magnitude. Thus human activities replenish the CO₂ deficit more than 45 times over and are equivalent to about 900 additional Mount Etnas.

Berner and Lasaga have characterized the calculation of CO₂ degassing from igneous and metamorphic activity as the most vexing problem encountered in modelling the carbon geochemical cycle¹⁰. In the hope of getting at least a reasonable approximation of CO₂ degassing over geological time, modellers have tied all degassing to seafloor

spreading rates. This approximation is indeed reasonable for CO₂ degassing at mid-oceanic ridges and subduction zones. The adequacy of seafloor spreading rates for predicting mid-plate volcano degassing rates is less clear, and it is possible that mid-plate volcano degassing is outside the conceptual framework of the current models. The high CO₂ degassing rates for Mount Etna underscore the need to ensure that mid-plate volcano degassing is satisfactorily represented in models of the carbon geochemical cycle. □

Terrence Gerlach is at the US Geological Survey, Cascades Volcano Observatory, 5400 MacArthur Boulevard, Vancouver, Washington 98661, USA.

THROMBIN RECEPTOR

Variations on a theme

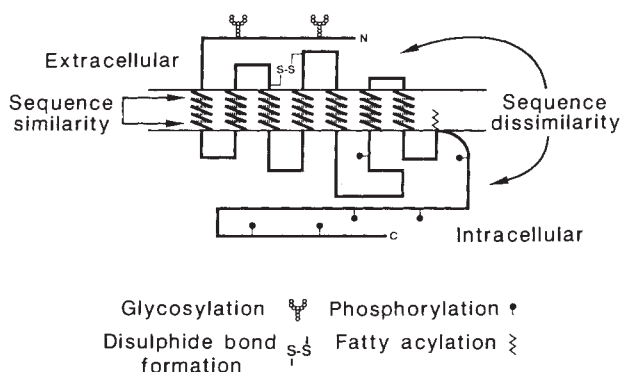
Robert J. Lefkowitz

HARDLY a week seems to pass without news of the cloning and sequencing of the complementary DNA for one or another G protein-coupled receptor with seven transmembrane domains — although each sequence is, of course, of interest, the structural and topographical features of these molecules have begun to seem, well, monotonous. But a report in *Cell* by Vu *et al.*¹, on the expression cloning of a thrombin receptor present on platelets and endothelial cells, now reveals a fresh variation on the structural theme and a unique mechanism of receptor activation. The deduced sequence has a putative cleavage site for the serine protease thrombin in the N-terminal extracellular domain of the receptor. The new terminus created by cleavage at this site apparently constitutes the physiological 'tethered' ligand for the receptor. The discovery will help to explain the molecular basis for one of the important haemostatic actions of throm-

bin, activation of platelet aggregation.

Of the various cellular pathways for transmembrane signalling, none has been more widely studied than that which involves specific cell-membrane receptors coupled by heterotrimeric, guanine regulatory proteins to effector enzymes and ion channels^{2,3}. The diversity of receptors which work in this way is remarkable. Moreover, virtually all of them share a highly conserved structure and topography² (see figure). The first member of this family whose sequence was elucidated was rhodopsin, the visual light 'receptor'^{4,5}. From hydrophobicity plots and a variety of physical measurements⁶, the presence of seven membrane-spanning domains was predicted, in analogy with bacteriorhodopsin, the distantly related bacterial proton pump for which electron diffraction information is available⁷.

Despite the obvious functional analogies between all G protein-coupled receptors, the publication of the deduced sequence of the mammalian β_2 -adrenergic receptor in 1986, which showed it had considerable sequence similarities and shared topographical features with rhodopsin, was greeted with some surprise⁸. As well as the seven transmembrane domains, these features include an extracellular N terminus with consensus sequences for N-linked glycosylation; sequence similarity, especially in the presumed membrane-spanning domains; sequence divergence in the connecting hydrophilic loops; a cytoplasmic carboxyl tail and/or large intracellular loop rich in serine and threonine residues, which are sites for regulatory



The polypeptide backbone of a prototypic G protein-coupled receptor crosses the plasma membrane seven times, with each membrane span assuming an α -helical conformation. Sites of post-translational modifications (glycosylation, phosphorylation, disulphide bond formation and fatty acylation) are shown, and amino-acid sequence conservation in the membrane-spanning domains and divergence in the hydrophilic connecting loops is indicated. Examples of ligands known to interact with the cloned receptors are listed in the box overleaf.

Strike one

ALTHOUGH notable for having afflicted public figures such as David Niven and Stephen Hawking, the cause of amyotrophic lateral sclerosis (also known as Lou Gehrig's disease) is still unknown. But about 10 per cent of all cases are familial, and one of the genes thought to be responsible has now been mapped. T. Siddique and colleagues (*New Engl. J. Med.* **324**, 1381–1384; 1991) examined 23 affected families, and demonstrated linkage in a subset of them using markers from the long arm of chromosome 21. The study points to the existence of at least one other disease locus, and marks an important step forward in characterizing the gene(s) involved and devising possible therapies.

Going up

SUPERCONDUCTING buckminsterfullerene is following a trend set by the perovskite superconductors four years ago: change an element or two and you get a massive increase in transition temperature. Only last month, A. F. Hebard *et al.* announced in *Nature* the totally unexpected finding that sheets of the crystallized C₆₀ balls became superconducting below a temperature of 18 K on doping with potassium. Now, writing in *Physical Review Letters* (**66**, 2830–2832; 1991), the same group reports that changing the dopant to rubidium raises the onset temperature to 28 K. And even this has been topped by K. Holczer and co-workers (*Science* **252**, 1154–1157; 1991), who find a transition temperature of 30 K, also with rubidium. Only (Ba,K)BiO₃ and the cuprate materials superconduct at higher temperatures.

Cancer connection

A POSSIBLE bacterial co-factor for gastric cancer has been identified by J. Parsonnet and colleagues (*Journal of the National Cancer Institute* **83**, 640–643; 1991). In patients with intestinal-type cancer attending two hospitals in California, the incidence of infection with *Helicobacter pylori* was at least 90 per cent, greatly exceeding the levels in the normal population. The study does not demonstrate that infection is causally related to the development of the cancers, but it seems likely that the bacteria may increase the chance of tumours arising (perhaps by stimulating production of cellular mutagens, or by inducing a high rate of cell proliferation following damage to the intestinal cells). Bacterial infections can often be treated successfully, so eradication of *H. pylori* may turn out to be a way of reducing the incidence of gastric cancer.

DIVERSITY OF THE KNOWN LIGANDS FOR CLONED G PROTEIN-COUPLED RECEPTORS

Biogenic amines	Epinephrine
	Norepinephrine
	Dopamine
	Acetylcholine
	Histamine
Tachykinins	5-Hydroxytryptamine
	Adenosine
	Substance P
	Substance K
	Neurokinin B
Glycoprotein hormones	Thyrotropin
	Follicle-stimulating hormone
	Luteinizing hormone
	Choriogonadotropin
	Parathyroid hormone
Polypeptide hormones	Angiotensin
	Arginine-vasopressin
	Vasoactive intestinal polypeptide
	Bombesin/gastrin-releasing hormone
	Thyrotropin-releasing hormone
Brain/gut peptide hormones	Thromboxane A ₂
	Light (i.e. retinal)
	Odorants
	Thrombin
	Endothelins
Arachidonic acid derivatives	Platelet-activating factor
	N-formyl peptide
	Mating factors (yeast)
	cAMP (<i>Dictyostelium</i>)
	Cannabinoids
Sensory stimuli	C ₅ A
Miscellaneous	

phosphorylation; and conserved sites for several other post-translational modifications, such as fatty acylation and disulphide bond formation (see figure). Despite the extraordinary diversity of ligands for these receptors (see table), these structural features have been found in most instances. The cloning⁹ of what seems to be a large superfamily of seven-transmembrane-domain receptors for odorants, which possess all of these features, greatly expands the size of the family. (Doron Lancet discussed these putative odorant receptors in detail in his News and Views article that appeared in last week's issue¹⁰.)

That seemingly minor variations on this structural arrangement have worked so well over millions of years of evolutionary time implies that the 'ground plan' represents a particularly efficient way to accomplish the job of transmitting an extracellular signal to the inner metabolic machinery of a cell. Nonetheless, our understanding of just how it works is still rudimentary. Ligands may interact with a 'pocket' created by some number of the (presumably) α -helical membrane-spanning regions, while other regions on the inner surface of the receptor interact with and activate specific G protein(s), thus determining the nature of the biological response³.

Even though the main structural and topographical features of the receptors with seven transmembrane segments have been highly conserved, several variations have been superimposed upon them to meet the demands of interacting with structurally

diverse ligands. In the case of small amine ligands, such as catecholamines, acetylcholine, histamine or serotonin, the membrane-spanning helices seem to form a pocket into which the ligand 'fits'. The same is probably true for the receptors for small peptides such as angiotensin and the tachykinins.

With the receptors for the big heterodimeric glycoprotein hormones, such as luteinizing hormone, thyroid-stimulating hormone and follicle-stimulating hormone, a large (over 300 amino-acid) N-terminal extracellular domain is present. This sugar-rich region contains the principal determinants for binding the hormones, and it may serve to position the large ligand to interact secondarily with the membrane-spanning regions. In the case of rhodopsin, a very specialized situation has evolved where the small chromophore retinal is covalently attached to the seventh transmembrane-spanning domain. Retinal nestles in a pocket created by the α -helices within the membrane and undergoes conformational changes when it absorbs a photon of light. In this instance, the light-modified retinal moiety in effect represents the 'ligand'.

Vu *et al.* describe a further highly unusual specialization of receptor structure in which the 'ligand' actually resides within the N terminus of the receptor molecule, and is only unmasked by a proteolytic clip mediated by thrombin. Interestingly, the site at which thrombin presumably cleaves the receptor is similar in sequence to that cleaved by thrombin in the zymogen protein C (LPDR/S or I, in the single-letter amino-acid code).

The biological effects of thrombin are not limited to aggregation of platelets, but probably include a variety of responses to vascular injury such as chemotaxis of monocytes, mitogenesis of lymphocytes and stimulation of production of platelet-derived growth factor by endothelial cells. It remains to be seen whether all of these responses are mediated by the same thrombin receptor, or whether a subfamily of such receptors exists (and, if so, whether they share the same activation mechanism). Beyond this is the fascinating issue of how many more variations on the theme of the seven-transmembrane-domain receptor remain to be discovered. □

Robert J. Lefkowitz is in the Howard Hughes Medical Institute Research Laboratories, Duke University Medical Center, Post Box 3821, Durham, North Carolina 27710, USA.

1. Vu, T. K. H., Hung, D. T., Wheaton, V. I. & Coughlin, S. R. *Cell* **64**, 1057-1068 (1991).
2. Gilman, A. G. *Rev. Biochem.* **56**, 615-649 (1987).
3. Dohlman, H. D., Thormer, J., Caron, M. G. & Lefkowitz, R. J. *Rev. Biochem.* **60**, 653-688 (1991).
4. Ovchinnikov, Y. A. *FEBS Lett.* **148**, 179-191 (1982).
5. Nathans, J. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4851-4855 (1984).
6. Applebury, M. L. & Hargrave, P. A. *Vision Res.* **26**, 1881-1895 (1987).
7. Henderson, R. *et al. J. molec. Biol.* **213**, 899-929 (1990).
8. Dixon, R. A. F. *et al. Nature* **321**, 75-79 (1986).
9. Buck, L. & Axel, R. *Cell* **65**, 175-187 (1991).
10. Lancet, D. *Nature* **351**, 275-276 (1991).

Safety in numbers

LAST week Daedalus was reflecting on the 'sperm-competition' which may go on in the reproductive tract of a woman with many recent lovers. He decided that sperm from different males must compete to fertilize the ovum by direct aggression — like fleets of rival battleships. He set DREADCO's biochemists to look for forms of chemical warfare between the sperm: toxins, anti-toxin defences, and so on.

He now muses that the battle may be more like a jamming contest between enemy radio transmitters. Sperm in the female reproductive tract do not wander vaguely around hoping to encounter an ovum somewhere. They swim purposefully towards it as soon as it appears. They seem attracted by a chemotactic effluvium of some kind. One obvious way for sperm to sabotage their rivals is to confuse them with false chemical signals.

This powerful form of competition has one serious problem. A jamming battle may simply have no winners. All parties may be utterly confounded by the mass of fake signals. A woman with sufficient lovers might thus escape pregnancy altogether. This, says Daedalus, may explain how prostitutes and courtesans managed in the days before reliable contraception. *Les grandes horizontales* relied on their many lovers to cancel each other out. In their scrutiny of sperm-battles under the microscope, DREADCO's biochemists are looking out for such 'suicide pact' effects. If they find them, a new and powerful sperm-jamming contraceptive may be possible.

But their real goal is subtler still. When the outcomes of sufficient sperm-battles are known, and their chemical weapons — toxins or false chemotactic emissions — are identified, it should be possible to divide sperm into specific classes, each vulnerable or invulnerable to known chemical weapons. Spermicides based on these substances will incapacitate the sperm of some men, but not those of others. They will form the basis of DREADCO's new range of selective spermicides.

These wonderful products will give women full control of their maternal ambitions. Some will be broad-brush formulations, child-proofing the user against whole classes of dishy undesirables, while leaving her open to more well-advised reproductive offers. Some will be more exactly targeted, and these will be individually compounded in DREADCO's Selective Fertility Clinic. One client may wish to be safely infertile towards her lover while reproductively open to her husband (or vice versa); another may wish to bear just one man's child from a galaxy of active admirers. The clinic will do its best to oblige. Clients may need to deploy a certain seductive cunning to obtain the relevant semen samples for analysis. David Jones

Will we reach the greenhouse?

SIR — Computer models indicate various degrees of temperature rise with a doubling of atmospheric CO₂ by about the middle of the coming century, based on continued growth in the rate of fossil fuel burning. (Current emission rate of carbon from human activities is about 6×10^9 tonnes per year, which would require approximately two centuries to double atmospheric CO₂.) These models assume that global growth in the CO₂ emission rate of the past two decades (between 1 and 2 per cent per year) will continue for the next 60 years. But will fossil fuels be available in sufficient quantity to permit such continued increase in CO₂ emission rate?

In the past century, food was produced with human and animal muscle power, and most US residents were farmers. Now, industrialized nations produce food with petroleum-fuelled mechanization, and only 2 per cent of the US population lives on farms. Mechanization permitted the huge increase in food production that has helped to foster unprecedented growth in human population. What fuel will support human societies beyond the year 2050, when the population will be more than twice that of today?

Global oil reserves cannot be limitless. We need to discover an additional $20\text{--}25 \times 10^9$ barrels a year over the long term. That is unlikely: estimates suggest that such discovery rates could be maintained for only several more years in the Middle East.

Reserves of natural gas in the United States have diminished more than 40 per cent during the past 20 years. Coal reserves, on the other hand, are voluminous. But if we ignore decline in natural gas production and project into the future an annual drop of 300,000 barrels per day in US oil production only, then the country's coal production rate will have to increase approximately 75,000 tonnes per year. This is certainly feasible, but if an increase in the rate of coal burning is to permit in addition a 1.5 per cent annual increase in total fuel burning to conform with greenhouse models, rate of coal burning will have to grow more than 6 per cent per year. This would reduce the reserve to a less than 40-year supply. The economically producible coal reserves of the rest of the world, although vast relative to cumulative production to date, are also too limited to allow a 1.5 per cent annual increase in total fuel burning for the next 60 years, if probable decline in global petroleum production during that interval is taken into account.

Thus, despite the importance of potential climate change and all its ramifications, we may nevertheless face a problem at least as intimidating as the greenhouse effect well before significant global warming has occurred. Climate change that decreases agricultural productivity of the world's primary food-producing regions becomes a secondary problem if we have largely exhausted the

fuel necessary to run the food-production system.

If we could return to the conservationist values of 1979–82, and make a permanent trend of diminishing petroleum consumption, the economic penalty would be painful, but we would extend our fuel-consumption time by many decades. Coupled with gradual, voluntary population decline, the human population could decrease until it is sustainable without the fossil-fuel conflagration of recent decades. This approach, unlike many others, does not challenge our technology to do something infeasible. But it constitutes a tremendous economic and moral challenge.

CRAIG BOND HATFIELD

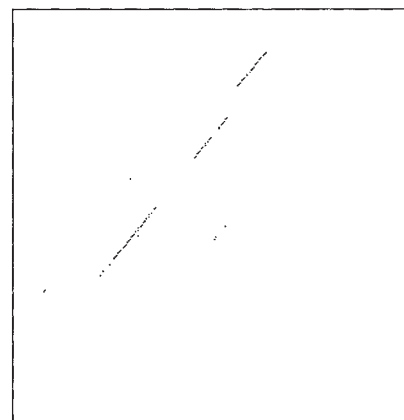
Department of Geology,
The University of Toledo,
Toledo, Ohio 43606, USA.

Drosophila Toll and IL-1 receptor

SIR — The generation of dorsoventral polarity in the early embryo of *Drosophila melanogaster* relies upon the expression of 12 maternal-effect genes termed the dorsal group. Mutations in any of these genes result in similar maternal-effect phenotypes in which all the cells of the blastoderm embryo develop along a path normally taken by the dorsal-most cells¹. Genetic studies have established the temporal sequence in which these genes act: *windbeutel*, *nudel*, *pipe*, *spatzle* and *gastrulation defective* are thought to act during oogenesis, whereas the other gene products are required zygotically in the order: *easter/snake*, *Toll*, *tube/pelle*, *dorsal*². Molecular analyses show that the products of the *snake* and *easter* genes are secreted serine proteinases^{3,4}, whereas that of *dorsal* is a transcription factor related to the avian viral oncogene *v-rel*⁵ and to the activator of immunoglobulin κ light-chain gene expression, NF κ B⁶. The function of the dorsal-group proteins results in the formation of a dorsoventral nuclear gradient of dorsal protein in the syncytial blastoderm stage embryo⁷. The gradient is generated by a process of differential nuclear localization and the phenotype of other dorsal-group genes is attributed to a failure of dorsal protein to accumulate in the nuclei^{8–10}.

The product of the *Toll* gene is a membrane protein consisting of extracellular transmembrane and cytoplasmic domains¹¹. The extracytoplasmic section of Toll protein has sequences in common with platelet glycoprotein 1b and other leucine-rich repeat proteins^{11,12}. Although Toll pro-

Interleukin-1 receptor



Toll

tein is uniformly localized in the egg plasma membrane of the syncytial embryo¹³, the observed gradient of nuclear accumulation by dorsal protein can be accounted for by specific activation of Toll receptor at ventral positions.

We have found that the cytoplasmic domain of Toll protein is related to that of the human interleukin-1 receptor (IL-1R)¹⁴. The identity extends for 135 amino acids throughout most of the domain and the significance of the highest scoring matches is extremely high, with identical or conservatively substituted amino acids being found at 45 per cent of positions (see Figs 1 and 2). It is thus probable that the biochemical nature of signal transduction by IL-1R and Toll protein is similar. Indeed, stimulation of cells by interleukin-1 causes the regulated nuclear localization of the transcriptional activator NF κ B, a homologue of the dorsal protein⁶. The inactive cytoplasmic form of NF κ B is complexed with a cytoplasmic anchoring protein, I κ B, and the phosphorylation of I κ B by protein kinase C causes it to dissociate, thereby enabling NF κ B to migrate to the nucleus¹⁵. Mutations in the *cactus* gene cause dorsal protein to accumulate in dorsal and

	418	428	438	448	458	460
IL-1R	QCGYKLFYIG	RDYVGEDIV	EVINENVKKS	RRLIILVRE	TSGFSWLGS	SEEQIATMYNA
	*	*	*	*	*	*
	*	*	*	*	*	*
	*	*	*	*	*	*
Toll	PQKPOLCVHE	RDWLVGCHP	ENIMRSVADS	RRTIIVLSON	FIKSEW--AR	LEFRAAHRSA
	885	895	905	915	925	935
	478	488	498	508	518	528
IL-1R	LVQDGKIVVL	LELEKIQDYE	KMPESINFIK	QKHGAIRWSG	DFTQGPQSGAS	TRFWKRVRYH
	*	*	*	*	*	*
	*	*	*	*	*	*
	*	*	*	*	*	*
Toll	LINEGRSRIIV	IYSDIGDVE	KLDEELKAYL	KMTIYKWDG	-----	PNFWDKLKFA
	945	955	965	975	985	995
	538	548				
IL-1R	MPVQRSPSS	KHQ				
	*	*				
	*	*				
	*	*				
Toll	LP-HRRPVGN	IGN				
	1005	1015				

FIG. 2 Alignment of Toll with the interleukin-1 receptor sequence. The sequences were aligned by reference to the DIAGON plot. Matches and conservative substitutions are indicated by asterisks.

lateral nuclei; the product of this gene is proposed to have an analogous function to $\text{I}\kappa\text{B}^8-10$.

At present the nature of IL-1R signal

1. Anderson, K. V. & Nusslein-Volhard, C. *Nature* **311**, 223-227 (1984).
2. Anderson, K. V. & Nusslein-Volhard, C. in *Gametogenesis and the Early Embryo* (ed. Gall, J.) 177-194 (Liss, New York, 1986).
3. DeLotto, R. & Spierer, P. *Nature* **323**, 688-692 (1986).
4. Chasan, R. & Anderson, K. V. *Cell* **56**, 391-400 (1989).
5. Steward, R. *Science* **238**, 692-694 (1987).
6. Ghosh, S. et al. *Cell* **62**, 1019-1029 (1990).
7. Steward, R. et al. *Cell* **55**, 487-495 (1988).
8. Rushlow, C. A. et al. *Cell* **59**, 1165-1178 (1989).
9. Roth, S., Stein, D. & Nusslein-Volhard, C. *Cell* **59**, 1189-1202 (1989).
10. Steward, R. *Cell* **59**, 1179-1188 (1989).
11. Hashimoto, C., Hudson, K. L. & Anderson, K. V. *Cell* **52**, 269-279 (1988).
12. Keith, F. J. & Gay, N. J. *EMBO J.* **9**, 4299-4305 (1990).
13. Anderson, K. V. in *Genes and Embryos* (eds Glover, D. M. & Hames, B. D.) 1-37 (IRL, Oxford, 1989).
14. Sims, J. E. et al. *Proc. natn Acad. Sci. U.S.A.* **86**, 8946-8950 (1989).
15. Ghosh, S. & Baltimore, D. *Nature* **344**, 678-683 (1990).
16. Mizel, S., O'Neill, L. A., Bird, T. A. & Saklatvala, J. *Immun. Today* **11**, 390-394 (1990).

transduction is a matter of debate. There is evidence that stimulation of IL-1R causes a rise in cyclic AMP levels and that the cAMP-dependent protein kinase A can activate the cytosolic form of $\text{NF}\kappa\text{B}$. Other work suggests that activation involves a GTP-binding protein and a serine-specific protein kinase distinct from protein kinases A and C^{16} . It has also been suggested that the cytoplasmic domain of IL-1R has a nucleotide binding activity and may act as an integral G protein¹⁴. In conclusion, the evidence now suggests a hypothesis of dorsal-protein nuclear localization that involves signal transduction by Toll receptor in a pathway similar to that mediated by IL-1R.

NICHOLAS J. GAY
FIONNA J. KEITH

Department of Biochemistry,
University of Cambridge,
Tennis Court Road,
Cambridge CB2 1QW, UK

Diffuse bands in emission

SIR — The origin of the diffuse interstellar absorption bands remains an unsolved problem after more than 50 years of research^{1,2}. Here I draw attention to a set of four emission features from the biconical nebula of the Red Rectangle³, the wavelengths of which correspond well with those of a recently classified subgroup of diffuse interstellar absorption bands⁴. It is thought that at least one of the diffuse band carriers is formed in the Red Rectangle region.

The diffuse interstellar band spectrum consists of over forty absorption features in the visible spectral region, with a wide variety of intensities and widths¹. On the basis of the relative strengths of a number of bands along different lines of sight, it has been suggested⁴ that they can be divided into families, one of which includes those bands occurring at wavelengths of 5,797, 5,850, 6,376, 6,379 and 6,614 Å. These bands are reasonably strong and comparatively narrow, their line widths being in the region of 1 Å. The Red Rectangle emission spectrum has an unidentified set of broader emission features with peak wavelengths of spatially integrated spectra near 5,799, 5,855, 6,380 and 6,615 Å (ref. 3) and typical widths of about 6 Å. These two sets of wavelengths are in remarkably good agreement and the greater widths and small wavelength shifts of the Red Rectangle bands relative to the diffuse interstellar bands can be attributed to the higher temperature close to the star HD44179. Also, the profiles of the bands in absorption (diffuse interstellar bands, DIB) and emission (Red Rectangle, RR) are qualitatively similar; the 5,797 (DIB⁵) and 5,799 (RR) profiles have a steep blue side, the 6,379 (DIB⁵) and 6,380 (RR) bands have approximately symmetrical profiles, and the 6,614 (DIB⁶) and 6,615(RR) bands also have a steep blue side. The notable absence

of other diffuse bands from the emission spectrum, such as those at wavelengths of 5,780 and 6,284 Å, does not necessarily mean that the carrier is not present; other reasons for non-observation, such as a low fluorescence quantum yield, may prevent detection.

The emission bands of the Red Rectangle are probably due to electronic transitions of one or more gas-phase molecules. The line widths and contours of the corresponding diffuse bands also lie within the known range of spectroscopic characteristics of cold large molecules. Considering briefly the suggestion that the C_{60} molecule or a derivative might be responsible for the diffuse bands⁷, it is worth noting that the magnitude of the rotational constant is nearly the same whether the C_{60} entity is neutral or ionized, or even has an included or surface atom. Preliminary simulations of the rotational contour of an electronic transition of a C_{60} entity at various temperatures show that a small decrease in rotational constant on electronic excitation is sufficient to cause the rotational contour to form a band head (steep blue side) and a red-degraded tail, as observed for the diffuse band/Red Rectangle 5,797/5,799-Å and 6,614/6,615-Å pairs, thus providing some support for the idea that a C_{60} entity might be responsible for the common set of bands observed in emission and absorption. Further work is in progress.

P. J. SARRE

Department of Chemistry,
University of Nottingham,
University Park,
Nottingham NG7 2RD, UK

1. Herbig, G. H. *Astrophys. J.* **196**, 129-160 (1975).
2. Bromage, G. E. Q. *J. R. astron. Soc.* **28**, 294-297 (1987).
3. Warren-Smith, R. F., Scarrott, S. M. & Murdin, P. *Nature* **292**, 317-319 (1981).
4. Krelowski, J. & Walker, G. A. H. *Astrophys. J.* **312**, 860 (1987).
5. Westerlund, B. E. & Krelowski, J. *Astron. Astrophys.* **203**, 134-144 (1988).
6. Herbig, G. H. & Soderblom, D. R. *Astrophys. J.* **252**, 610 (1982).
7. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162-163 (1985).

Stress and evolution

SIR — According to Sheldon¹, gradual phyletic evolution can be sustained only by organisms living in or able to track narrowly fluctuating, slowly changing environments, whereas stasis, almost paradoxically, seems to prevail in widely fluctuating, rapidly changing environments. This generalization, expanded in a recent model², relates in a geological time scale to physical variables such as sea level, substrate and climate. I argue here that it is also consistent with evidence from evolutionary genetics.

Organisms first encountering widely fluctuating, rapidly changing environments would be characterized by high stress levels. Assuming stress to be an environmental factor that causes a potentially injurious change in a biological system³, its main impact is at the extreme end of the stress gradient, for example at the boundaries of species, where fitness becomes limiting because of elevated energetic costs. Although geographical limits depend on many different factors, the energetic consequences of climatic extremes are often of major importance^{3,4}. The resistance of organisms to a wide range of stresses can be increased by artificial selection, but at the cost of lowered metabolic rate and hence fitness, which would be restrictive for evolutionary change⁴.

Severe environmental stress increases the phenotypic and genotypic variability of ecologically important traits, including those involved in stress resistance^{3,4}. This would be expected at species boundaries, where environmental fluctuations tend to be too severe for adaptation to the stresses. Even with such high variability, the concomitant metabolic and fitness costs would preclude expansions of the species³ range into ecologically more extreme habitats, so that rapid evolutionary change would be unlikely. The opposite extreme is an invariant and nonstressful habitat, such as may occur under domestication, where low genetic variability for stress-resistance traits³ is restrictive for substantial change despite the availability of sufficient metabolic energy. (In assessing genetic variability, stress-resistance traits are considered directly rather than by the common approach of looking at electrophoretic variants, which cannot provide a complete understanding of traits at the phenotypic level^{3,4}.)

Evolutionary change may therefore be most rapid in populations from habitats intermediate between these extremes, where there is some environmental stress associated with sufficient metabolic energy to allow some adaptive change. Furthermore, high phenotypic plasticity is likely in such habitats³. Also, ecological systems may be the most diverse under a certain degree of stress⁵, so that the greatest number of species occurs when disturbance is intermediate⁶. These circumstances may be a feature of the

narrowly fluctuating tropics, where the highest speciation rates are likely². Indeed, the diversity of species in the genus *Drosophila* in Australia is highest in the humid tropics, whereas at more widely fluctuating heat-stressed forest edges, species diversity is lower, corresponding to the depauperate faunas characteristic of the heat-stressed habitats of the wet-dry tropics⁷.

In conclusion, and in agreement with the palaeontological evidence^{1,2}, maximum speciation rates are expected in habitats characterized by narrowly fluctuating environments implying moderate stress, moderate genetic variability of ecologically important traits, and not unduly restrictive metabolic costs. In this way, gaps between palaeonto-

logy, physiology, ecology and evolutionary genetics are being bridged with the level of physical stress, especially climatic, being the connecting environmental link.

P. A. PARSONS

Department of Zoology,
University of Adelaide,
Adelaide, South Australia 5000,
Australia

1. Sheldon, P. R. *Nature* **330**, 561–563 (1987).
2. Sheldon, P. R. *Nature* **345**, 772 (1990).
3. Hoffmann, A. A. & Parsons, P. A. *Evolutionary Genetics and Environmental Stress* (Clarendon, Oxford 1991).
4. Parsons, P. A. *Trends Ecol. Evol.* **5**, 315–317 (1990).
5. Moore, P. D. *Nature* **306**, 17 (1983).
6. Petraitis, P., Latham, R. E. & Neisenbaum, R. F. *Quart. Rev. Biol.* **64**, 393–418 (1989).
7. Parsons, P. A. *Proc. ecol. Soc. Aust.* **13**, 43–49 (1985).

Schroedinger's quantum cat

SIR — Schroedinger's imaginary cat^{1–3}, supposed to be in a linear combination of the quantum states live and dead, emphasizes the contrast between classical and quantum systems — and of the ambiguity of the boundary between them in the usual interpretation of measurement in quantum mechanics. Dissatisfaction with such ambiguity has led many physicists to construct alternative quantum theories. It may be held, of course, that the states of cats, which are macroscopic systems, are irrelevant to quantum measurement, but is that necessarily so?

Consider a molecule of DNA, which can be mutated genetically by the influence of photons. The orthodox view⁴ is that a quantum measurement is "any process of interaction between classical and quantum objects occurring apart from and independently of any observer". It is important to be clear about the distinction between the classical and quantum systems used in measurement. The states of classical systems are robust and can be preserved and reproduced, with only small probability of error, like the characters on this page or the state of survival of a real cat. They form a classical record.

By contrast, the reproducibility of quantum systems is constrained by the no-cloning theorem of Wothers and Zurek⁵ and Ghirardi and Weber⁶. The information in a set of quantum states can easily be destroyed.

It is usually assumed that the classical measuring apparatus must be physically macroscopic, but this is not the case for a molecule of DNA. Yet the sequence of base pairs is a classical record, cloning is the norm, parts of the record may be retained over millions of years and may be reproduced more often than any published record of a measurement. With the help of biological proof-reading, the error in copying DNA is estimated to be between 10⁻⁷ and 10⁻¹¹ per replication of a base pair⁷.

But DNA also responds to quantum events, as when mutations are produced by single photons, with consequences that may be macroscopic — leukaemia, for example.

Moreover, these consequences may be much delayed, as in recessive mutations, so that a classical DNA record may be kept at the molecular level for several generations before it is expressed macroscopically. Such a mutation is a classical record, on the molecular scale, of the quantum event that produced the mutating photon, so that it represents a quantum measurement. Thus quantum measurements do not require macroscopic systems. The total energy of the

bonds that bind a base pair can be less than the energy of the photon that produces the mutation, so that quantum measurements do not require amplification either.

Schroedinger chose his example well. Most people find it difficult or impossible to believe that an object of affection such as a cat could be in a state which is a linear superposition of a live state and a dead state. But if a classical record is what matters, a quantum superposition of viable and lethal sequence of base pairs in DNA is as absurd as a cat that is both alive and dead. Should one seek to learn more about Schroedinger's cat by studying her kittens?

IAN PERCIVAL

School of Mathematical Sciences,
Queen Mary and Westfield College,
Mile End Road, London E1 4NS, UK

1. Schroedinger, E. *Naturwissenschaften* **23**, 807–812 (1935).
2. Schroedinger, E. *Naturwissenschaften* **23**, 823–828 (1935).
3. Schroedinger, E. *Naturwissenschaften* **23**, 844–849 (1935).
4. Landau, L. & Lifshitz, E. *Quantum mechanics* (Pergamon, London, 1958).
5. Wothers, W. K. & Zurek, W. H. *Nature* **299**, 802–803 (1982).
6. Ghirardi, G. C. & Weber, T. *Nuovo Cim.* **78B**, 9–20 (1983).
7. Watson, J. D. et al. in *Molecular Biology of the Gene* (Cummings, Menlo Park, 1988).

Whose larvae?

SIR — The inability to distinguish between marine invertebrate larvae of closely related taxa is a long-standing problem in marine ecology. Within groups such as bivalves or barnacles, the larvae of many species are so similar that they cannot be identified by morphology alone¹. Using the polymerase chain reaction to amplify and sequence mitochondrial DNA, we are able to distinguish between the morphologically identical, co-occurring larvae of two species of sea cucumber. This technique should prove effective for the identification of larvae even when they are far from their adult population, such as open-ocean (teleplanic) larvae or the larvae of deep-sea organisms.

The sea cucumbers *Psolus fabricii* and *Cucumaria frondosa* are common members of the subtidal communities along the north-east coast of the United States. *P. fabricii* occurs primarily on hard surfaces, *C. frondosa* is found on both hard-rock and soft-sediment bottoms^{2,3}. In 1990, we con-

Psolus ACT-GAAACAGATATAGTTACCGCAG

Cucumaria C..G.T.....A.A.....

Psolus GGATAACAGCGTCATCTCCTTTAAGAG

CucumariaA.....

Sequence of a portion of the 16S rRNA gene from two species of sea cucumber in the Gulf of Maine. Individual larvae were placed directly into an amplification buffer containing primers to the conserved region of the gene. Sequences were determined directly from the amplified products using a nested pair of primers.

ducted plankton tows around the Isles of Shoals, New Hampshire, to collect larvae of *P. fabricii*. Bright red pentacula larvae were caught in large numbers in March, and then again in late May. Based on the advice of local embryologists, and on the similarity of the larvae to those of the west coast species *Psolus chitinoideus*, we assumed them to be from the locally abundant species *P. fabricii*. Furthermore, their red pigmentation matched the colour of adult *P. fabricii*.

As part of a population study, we began sequencing a portion of the 16S ribosomal RNA gene from adult *P. fabricii* and the collected larvae. To our surprise, the nucleotide sequence of the larvae differed from that of the adult *P. fabricii* by 17 per cent over 389 base pairs (see figure), which is a substantially greater difference than is usually observed between conspecific echinoderms⁴. Further work showed that the larval sequence matched that of adult *C. frondosa* exactly. In addition, the genetic analysis told us that *P. fabricii* larvae were also present in at least one of the plankton samples, though we were not able to recognize them by morphology alone. Using the sequence differences between these two species, we are now in a position to create genetic probes that can be used to identify hundreds of larvae at a time.

The use of the polymerase chain reaction for identifying marine invertebrate larvae opens up a wide range of possibilities for field studies of dispersal and recruitment. For example, isotopic studies of deep-sea gastropod shells have shown that their larvae must ascend to the near-surface waters^{5,6}, but no one has ever collected a larva from this region that could be shown conclusively to be a deep-sea species. With the establishment

of libraries of mitochondrial DNA sequences, the problem of identifying morphologically similar invertebrate larvae will at last be solved.

RICHARD RANDOLPH OLSON
JONATHAN A. RUNSTADLER
THOMAS D. KOCHER

Department of Zoology and
Center for Marine Biology,
University of New Hampshire,
Durham, New Hampshire 03824, USA

1. Smith, D. L. *A Guide to Marine Coastal Plankton* (Kendall Hunt, Dubuque, Iowa, 1977).
2. Pawson, D. L. *N.O.A.A. Tech. Rep.* **405** (Houghton Mifflin, Boston, 1979).
3. Gosner, K. *Field Guide to the Atlantic Seashore* (Houghton Mifflin, Boston, 1979).
4. Palumbi, S. & Wilson, A. C. *Evolution* **44**, 403-415 (1990).
5. Killingly, J. S. & Rex, M. A. *Deep Sea Res.* **32**, 809-818 (1985).

Gallo's virus sequence

SIR — In our recent Scientific Correspondence¹, we provided evidence that virus stocks received at our laboratory in 1983 from the Institute Pasteur contained mostly species similar to each other but quantitatively different from published sequences of HTLV-IIIB and LAV-1 (refs 2,3). After depositing the sequences in the Los Alamos database, we learned of an interview⁴ in which Dr G. Myers, curator of that database, said that our published sequences contain typographical errors compared with the deposited sequences as well as with the published sequence of LAV-1 (ref. 3). These errors were introduced during transcription of the figures by our graphics artist. The correct sequences were deposited in the Los Alamos database, and these and related sequences have also been deposited at EMBL with the accession numbers X57446 to X57466. The correct sequences are slightly more divergent than the previously published sequences.

We have also learned that our method for computing percentage similarity was considered to have given an "inflated" impression of how different the viruses are because of our handling of gaps in sequences. There are four possible approaches for handling gaps in computer-aligned DNA sequence similarities⁵⁻⁸: (1) ignore the gaps in estimating sequence divergence; (2) count each gap as equivalent to a base substitution; (3) weight each gap with an arbitrary penalty to account for a perceived rareness of deletion/duplication events relative to nucleotide sub-

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

SIMILARITY OF LAV-1/HTLV-IIIB SEQUENCES

	1985 Published sequences of		Analysis in 1990 of samples received in 1983		
	HTLV-IIIB	LAV-1	No. 1	No. 2	No. 3
HTLV-IIIB	—				
LAV-1	98.5 (98.1)	—			
JBB sample No. 1 ¹	93.6 (91.3)	94.1 (91.5)	—		
JBB sample No. 2 ^{1*}	93.9 (91.5)	94.3 (91.9)	99.0 (99.0)	—	
JBB sample No. 3 ¹	94.0 (90.5)	93.8 (90.6)	98.9 (98.9)	98.4 (98.4)	—

Percentage similarities were calculated by using each gap introduced as a single nucleotide change. Numbers in parentheses refer to the values previously obtained. HTLV-IIIB data from ref. 2; LAV-1 data from ref. 3.

*The JBB samples were 870-875 base pairs long except for sample 2, which was 670 base pairs long.

stitutions; and (4) count each base within a gap as a mismatch. The computer algorithm we used was the Data Bank Search function in the MicroGenie DNA analysis software package (Beckman Instruments), which uses the fourth method. We now realize that methods 1 or 2 would have been more appropriate, as they are more conservative and more commonly used in estimating sequence divergence, and we regret the confusion that the earlier calculations created.

Recalculated values of our table using method 2 are presented here and these results (see table) still suggest that the majority of species recovered from the three 1983 Institute Pasteur samples we analysed are more different from the LAV-1 and HTLV-IIIB strain sequences (all available in the Los Alamos database) than they are from each other. We look forward to developments in studies in natural histories of virus populations as well as to sequences from

other early isolates, including LAV-1 in Paris, that may extend our understanding of the extent and type of variability that exists in these important disease agents.

MARVIN S. REITZ, JR
HOWARD Z. STREICHER
ROBERT C. GALLO

Laboratory of Tumor Cell Biology,
National Institutes of Health,
National Cancer Institute,
Bethesda, Maryland 20892, USA

1. Guo, H.-G. *et al. Nature* **349**, 745-746 (1991).
2. Ratner, L. *et al. Nature* **313**, 277-284 (1985).
3. Wain-Hobson, S. *et al. Cell* **40**, 9-17 (1985).
4. Palca, J. *Science* **251**, 1306 (1991).
5. Li, W.-H. & Graur, D. *Fundamentals of Molecular Evolution* (Sinauer, Sunderland, Massachusetts, 1991).
6. von Heijne, G. *Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit* (Academic, New York, 1987).
7. Weir, B.S. (ed.) *Statistical Analysis of DNA Sequence Data* (Dekker, New York, 1983).
8. Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure Vol. 5, Suppl. 3* (National Biomedical Research Foundation, Washington, DC, 1982).

... and his response

SIR — The nucleotide sequence similarity of LAV-BRU and HTLV-IIIB, the two main early isolates of HIV, has been a subject of much controversy. New work shows that in 1983 at the Institute Pasteur a virus isolated from patient BRU (originally called LAV-BRU, then HIV-BRU, and now HIV-JBB) was unknowingly contaminated with virus from another patient, LAI (virus now called HIV-LAI)¹⁻³. HIV-LAI seems to grow extremely well in cell culture, and reports from other laboratories show that HIV-LAI frequently contaminates cultures of viruses from people with AIDS (see ref. 2 for references). Samples of virus thought to be LAV-BRU were sent in 1983 from the Institute Pasteur to the National Institutes of Health (NIH). In some cases these samples contained only HIV-JBB, but in one case the sample also contained HIV-LAI.

It also appears that cultures of virus from people with AIDS became contaminated with HIV-LAI at NIH. When workers at the Institute Pasteur and at NIH later cloned viruses from AIDS patients, both apparently sequenced HIV-LAI. At the Institute Pas-

teur this virus was designated LAV-BRU. At NIH this virus was designated HTLV-IIIB.

Some interesting epidemiological questions remain about the origin of HIV-LAI and its relationship to other HIV strains from early in the AIDS epidemic. But the similarity of the nucleotide sequences of LAV-BRU and HTLV-IIIB now seems to be explained. I note that both laboratories had other isolates in 1983-84 and, in short, none of this affects the history of the important events written for *Nature* in 1987⁴ and *Scientific American* in 1988⁵ by L. Montagnier and myself. It is now time for this period of controversy to come to an end and for us all to focus our efforts on ending the pandemic.

ROBERT C. GALLO
Laboratory of Tumor Cell Biology,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 20892, USA

1. Guo, H.-G. *et al. Nature* **349**, 745-746 (1991).
2. Wain-Hobson, S. *et al. Science* **252**, 961-965 (1991).
3. Chermann, J.-C. *et al. Nature* **351**, 277-278 (1991).
4. Gallo, R.C. & Montagnier, L. *Nature* **326**, 435-436 (1987).
5. Gallo, R.C. & Montagnier, L. *Scient. Am.* **259**, 41-48 (1988).

This strange Universe

John Peacock

Basics of Modern Cosmology. By A. D. Dolgov, M. V. Sazhin and Ya. B. Zeldovich. Editions Frontières: 1991. Pp.247. Hbk, \$40; pbk, \$28.

MORE than three years after his death, a new book has appeared bearing the name of one of this century's greatest physicists. Yakov Borisovich Zeldovich (1914–1987) was a man of almost unequalled breadth of interests, who made fundamental advances in gas dynamics, particle physics and cosmology. Many who never had the privilege of meeting him (including myself, sadly) have got to know the Zeldovich style through his monumental two-volume works on gas dynamics (with Y. P. Raizer) and relativistic astrophysics (with I. D. Novikov), together with innumerable research papers. These writings are characterized by an admirable directness: Zeldovich usually goes straight to the heart of a problem, rarely getting distracted in the erection of unnecessary formalism. Some of his most influential papers are almost unbelievably short. The impression that emerges is of a scientist equipped with great depth of insight, and not someone likely to have had much patience with the bureaucratic society in which he lived and worked. Some of these features emerge from the obituary by Andrei Sakharov (see *Nature* **331**, 671; 1988), but not the story I like best: of Zeldovich cannily protecting himself from police harassment by donning his medals before going out and getting drunk.

Basics of Modern Cosmology thus has a rather high standard to live up to, and I have to say I found it a little disappointing by comparison with the expectations it raised. It is a translation of a book originally completed in 1986, and based upon lectures given by Zeldovich, although more than two-thirds of the text was written by A. D. Dolgov and M. V. Sazhin. The aim is to give an introduction to the revolution in cosmology that occurred in the first half of the 1980s, caused by the ingestion of methods and ideas from particle physics. The book can be divided into four parts: (1) principles of pre-inflation cosmology; (2) a primer on particle physics; (3) a discussion of scalar fields and inflation, including baryosynthesis and generation of fluctuations; (4) the large-scale structure of the Universe. This is an awful lot of material to fit into 247 pages, especially given that the book tries to be very complete. The result is, inevitably, that most topics are treated rather briefly. This is a pity, given that the amount of introductory material indicates that the book is probably intended for students unfamiliar with modern cosmology. Such

readers will be frustrated by the many formulae that are pulled out of the air without derivation or reference, particularly where these relate to items of technical difficulty very much greater than the surrounding material (such as the casual use of the path integral over euclidean metrics in chapter 9). Novices in non-Soviet countries will also be

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Zeldovich — equipped with great insight.

hampered by the fact that the terminology often fails to connect with that common in Western research literature. Nothing wrong with that in principle: the fact that most of the book's references are to Soviet literature is a valuable antidote to the all-too-common belief that all new developments in cosmology appear in the *Astrophysical Journal*. But this approach has been carried too far if it means that the spectrum referred to in the West as 'Zeldovich', ' $n = 1$ ' or 'scale-invariant' (one of Zeldovich's most important bequests to cosmology) is apparently not mentioned, unless you know what you are looking for. It is quite possible to give the Soviet view of the subject without being obscure, or using a widely varying level of difficulty, as is shown in the recent excellent text by A. Linde (*Particle Physics and Inflationary Cosmology*, Harwood, 1990).

So, if this is not the place to send an ill-equipped graduate student, what will the established professional researcher find here? Several aspects of the book bear the strong imprint of Zeldovich's views on the subject, which are sometimes at variance with Western orthodoxy, and make for thought-provoking reading. Zeldovich always placed great weight on the fact that the energy of a closed Universe is zero; the idea that the 'spontaneous birth' of a closed Universe does not violate energy conservation and therefore might be preferred in a quantum sense is treated quite extensively. Also, a whole chapter is devoted to A. A. Starobinsky's ideas on inflation being driven by vacuum polarization (that is, quantum corrections to the energy-momentum tensor) rather than by a scalar field. This alternative is subsequently dropped without explanation, however, and the book then proceeds along more standard lines. This is a pity, as I feel an opportunity was lost for taking a more critical look at some of our current ideas. In particular, the crucial defect of inflationary cosmology as it stands is that the scalar field must be very weakly coupled, to avoid the overproduction of inhomogeneities. Unlike Guth's original proposal, its existence is thus not at all well motivated from particle physics. This weakness deserves to be seized on, especially by a book that is not afraid of being a little heterodox, but it passes almost without comment.

The last two chapters are on fluctuations and large-scale structure, and I particularly looked forward to seeing Zeldovich's last views on a field he helped to create. In the end I was disappointed, especially because very little is said about confrontation with observations. Perhaps the authors consider this to be unnecessary or uninteresting: "The formation of the large-scale structure of the Universe is already understood qualitatively and only the quantitative details need to be worked out." This may well be true in the sense that brain surgery is only applied quantum mechanics, but it does give a rather peculiar view of the subject. One of the most exciting aspects of cosmology in recent years has been the rapid accumulation of evidence about the structure of the Universe we inhabit. Different researchers will disagree about whether this process is converging to any sort of understanding, but it is undeniably a source of great intellectual enjoyment. A book on this subject, especially one for students, should be able to communicate the sheer excitement of being able to understand for the first time the main features of this strange Universe. Sadly, this one does not. □

John Peacock is at the Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK.

On the rocks

J. Oerlemans

The Growth and Decay of Ice. By G. S. H. Lock. Cambridge University Press: 1990. Pp.434. £65, \$100.

WRITING a treatise on the appearance of ice on Earth on all possible space scales, from the molecular to continent-sized ice sheets, is truly ambitious. From a general scientific point of view, I can appreciate the charm of placing our knowledge of ice in a single framework, naturally built around the basic physical principles involved.

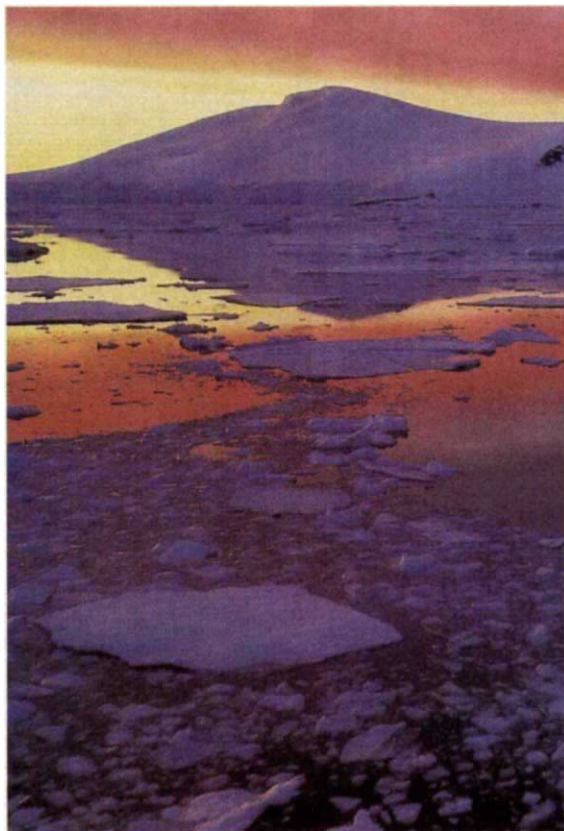
When reading *The Growth and Decay of Ice*, my initial enthusiasm turned into doubt. In which sense, for instance, can freezing in an ice-resistant insect be related to the effect of solar insolation variations on ice sheets? This book actually shows that parallels in conceptual development of methodology between the study of ice on the molecular/microscale and the study of its role in the large-scale geophysical environment are hard to find. Frozen water is the binding element, but that is all.

After a general introduction, two chapters are devoted to a thorough treatment of the thermodynamics of ice. This treatment is formal and rigorous, and of considerable didactic quality. Then chapters follow on 'ice and water', 'ice and air', 'ice and Earth', 'ice and life' and, finally, 'decay of ice'. It is not surprising that these topics fill over 400 pages. With respect to 'ice and water', an interesting discussion is provided on the effect of convection (free and forced) on planar growth, and a short but instructive view is given on freezing on submerged bodies. In 'ice and air' the discussion centres around the formation of cloud droplets and ice crystals, riming and icing on off-shore and land-based structures. Surface energy exchange is the starting point in the chapter on 'ice and Earth', followed by sections on the pores system, with reference to the surface morphology of glaciers, and on permafrost. In 'ice and life', topics of study are freezing in biofluids, small organisms, plants, animals and, partly as an interesting application, foods. The final chapter on the 'decay of ice' is surprisingly short.

Throughout the book, the treatment loses considerable depth when larger-scale features are considered in the global context. Many statements on glaciations, ice ages, climate sensitivity, large-scale snow cover

and sea ice are superficial and based on old material (for instance, a single sentence stating that S. H. Schneider and T. Gal-Chen have calculated (in 1973) that a two per cent drop in the solar constant would lead to a global temperature below 0 °C, without further discussion). For a text of this type, the number of listed papers from the past decade is remarkably small. This cursory treatment is not justified; in the past 10–20 years, a tremendous amount of work has been done on such topics. Furthermore, the possibilities offered by satellites and powerful computers are hardly mentioned.

In making a personal judgement, it appears natural to split the material in this book into two parts. The first is on the



Sunset on ice – the Lemaire Channel in the Antarctic (taken from *Wild Ice*, by R. Naveen, C. Monteath, T. De Roy and M. Jones (Smithsonian Institution, 1990, \$29.95)).

thermodynamics of ice, nucleation, growing of crystals under various conditions, that is, the microscale aspects. The treatment of these subjects is of high quality and studying it can be advised to anyone involved in ice-related projects. The second part then deals with the macroscale aspects, a bit scattered through the book (mainly chapter 1 and parts of chapters 6 and 8). This material provides a superficial introduction to the role of ice in regional and global systems, which is not up to date. □

J. Oerlemans is at the University of Utrecht, Institute of Meteorology and Oceanography, Princetonplein 5, 3584 CC, Utrecht, The Netherlands.

Charting a revolution

Kevin Struhl

DNA Science: A First Course in Recombinant DNA Technology. By David Micklos and Greg Freyer. Cold Spring Harbor Laboratory Press: 1990. Pp.477. \$29.95.

In the early days of the recombinant-DNA revolution, it was frequently remarked that experiments were so easy that they could be performed by high-school students. It is now nearly 20 years later, and despite extensive publicity and controversy, there are few high schoolers who are familiar with the theory or practice of recombinant-DNA technology. Even at the college level, courses in this subject are generally restricted to the small minority of upper-level undergraduates in biology. Given that both recent and future advances in biology will undoubtedly have profound effects on society, the current level of ignorance is deplorable.

Although such ignorance has many causes, one problem is the lack of suitable books that convey the excitement, power, knowledge and actual techniques of recombinant-DNA technology. There are, of course, several molecular-biology cookbooks that are designed to instruct 'non-experts' in the art of performing the vast array of available experimental techniques. But these collections of current protocols are intended for laboratory researchers, not high-school or college students. Thus, there is a crucial need for a book that introduces true novices to basic experimental concepts and techniques, although maintaining high scientific standards and relevance.

This niche is now admirably filled by *DNA Science: A First Course in Recombinant DNA Technology*, by David A. Micklos and Greg A. Freyer. The combined skills and experience of the authors, one a science educator and the other a practising molecular biologist, have resulted in a unique book. It has elements of a textbook, laboratory manual, historical novel and popular magazine article, yet is remarkably cohesive. In particular, there are numerous connections between the scientific concepts advanced in the first half of the book and the laboratory experiments that constitute the second half.

The text is written in a journalistic style that can be read from cover to cover (but not in one sitting). The two initial chapters cover some basic themes of molecular biology by tracing the history of DNA science from Linnaean classification of species through the cracking of the genetic code. By emphasizing what the important questions were and how they were experimentally solved, the science is lucidly explained while conveying the excitement of discovery. The pivotal third

chapter blends the historical development of the recombinant-DNA revolution (starting with the restricting-modification phenomenon) with a reasonably detailed description of the basic tools and techniques of the trade. The historical account is generally accurate, but does suffer from the usual problem of neglecting the critical contribution of Peter Lobban. The real value of this chapter, however, is that it provides an excellent theoretical underpinning for the laboratory experiments that follow.

Chapter 4 describes more advanced experimental techniques such as generating and screening genomic and complementary-DNA expression libraries, monoclonal antibodies, DNA sequencing, Southern and Northern. By necessity, the coverage of these topics is superficial, which sometimes has the effect that technical details follow too quickly from the initial description of the experimental approach, thus leading to confusion on the part of the novice. The final four chapters of the text concern the application of recombinant-DNA technology to studies of gene regulation in development and cancer as well as applications to human genetics, agriculture, medicine and industry. The topics represent an impressive range of current problems and techniques and are generally well described. Some minor criticisms are the inadequate formulation of the crucial questions in developmental biology and the treatment of the polymerase chain reaction. Nevertheless, the quality of the textual part of the book is excellent.

The laboratory experiments are based on the authors' experiences at the educationally innovative DNA Learning Center at the Cold Spring Harbor Laboratory, and include making measurements, sterile technique, DNA-restriction analysis, construction of recombinant DNA molecules, transformation of *Escherichia coli*, simple genetic techniques such as replica plating, and purification and identification of recombinant DNA. These are excellent choices because they can be performed in unsophisticated laboratory settings, yet are fundamental experiments of modern technology; the description of the actual protocols is outstanding. Among the most important features are the extensive prelaboratory notes and preparations, the diagrammatic representation of the experimental manipulations at the beginning of each section, the detailed explanations of the individual steps, the field guide to electrophoresis effects and the questions and future research directions discussed at the end of each experiment. On a more practical note, the authors have collaborated with the Carolina Biological Supply Company so that instructors can order all the necessary reagents and equipment from a single source. □

Kevin Struhl is in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA.

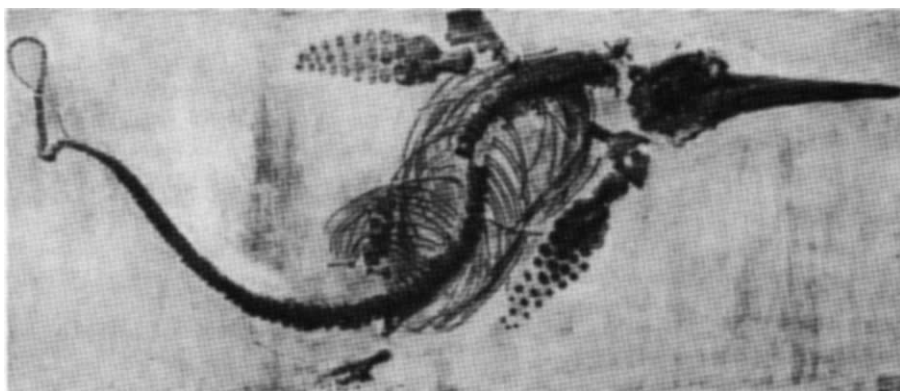
Dances with dinosaurs

Henry Gee

Dinosaurs, Spitfires, & Sea Dragons.

By Christopher McGowan. Harvard University Press: 1991. Pp.365. \$29.95, £23.95.

My favourite film is Walt Disney's *Fantasia*, in which a troupe of hippopotami cavort in balletic formation. The spectacle of dancing hippopotami is, of course, ridiculous, but therein lies its charm. The spectacle of dancing dinosaurs would seem equally risible but for the fact that many contemporary restorations, depicting dinosaurs as every bit as



One of the most complete skeletons of *Leptopterygius tenuirostris* (originally referred to as *Ichthyosaurus tenuirostris*) from Street, Somerset, England.

active as modern mammals and birds, are ultimately based on dinosaur fossils and what can be inferred therefrom. The problem is that the chain of reasoning between solid research and the more frivolous end of the pop-dinosaur market may be worn extremely fine. In which case dinosaurs become indistinguishable from cartoon characters.

In *Dinosaurs, Spitfires & Sea Dragons*, Christopher McGowan aims to set the record straight by setting out to discover exactly what can be said about dinosaurs on the basis of the evidence, and what cannot. Dinosaur colour may be debated endlessly for the lack of hard data, but it is hard to take issue with the laws of physics that govern locomotion, whether of dinosaurs, dromedaries or ducks. As with R. McNeill Alexander's similar book *Dinosaurs and Other Extinct Giants* (Columbia University Press), chapters start with brisk physics lessons in mechanics (both statics and dynamics) and go into fluid dynamics, acoustics and the principles of aeronautics where necessary. McGowan then applies these basic principles to a range of dinosaur-sized problems that have furrowed palaeontological brows for decades.

He does not debunk 'dancing' dinosaurs, the idea of endothermy in dinosaurs or any other fashionable contention — he seeks only to appraise the problem in the light of

simple physics and what we know about living animals. This may itself be surprisingly little. Zebras, for example, "are remarkably uncooperative at running in straight lines for biologists to clock their speeds", he says. With fine sensitivity to the diversity of life's solutions to physical problems, he concludes (for homeothermy) that the probability of some dinosaurs being homeothermic is greater than for others — reflecting the diversity we see today. Even mammals, for example, are not universally homeothermic to the same degree. Monotremes and some insectivores have remarkably low basal metabolic rates, and temperatures can fluctuate depending on whether the animal hibernates or not. Nobody would seriously posit homeothermy in extant animals as a problem with a clear-cut answer: the same was undoubtedly true for dinosaurs, a

diverse and long-lived group.

In a similar way, dinosaurs may well have danced, but each species followed its own particular choreography. Even if they were 'cold-blooded' like modern reptiles, there is no reason to suppose that small, lightly built theropods could not have put on a fair turn of speed. Anyone who has ever tried to chase a gecko across the ceiling will testify to that. High mammalian or avian metabolic rates, then, are not prerequisites for high levels of activity in reptiles, provided they can rest after bursts of activity to work off an oxygen debt accrued by anaerobic glycolysis. Physique, too, dictated that large sauropods would have been eternal wallflowers, condemned to a slow amble that nevertheless, because of their enormous size and the effects of scale, would have been faster than that achievable by many human runners.

The appliance of science to dinosaurs shows that one can, in fact, say quite a lot without launching oneself into Disneyland. McGowan does the same thing for pterosaurs and flight, and allows himself a chapter on ichthyosaurs that seems rather lop-sided in that he discusses the known genera, their characteristics, stratigraphic ranges and excavation history in addition to their mechanics. This detail is absent in his discussion of dinosaurs and pterosaurs, but he may be excused this indulgence given that he

is an ichthyosaur specialist.

For all its physics, *Dinosaurs, Spitfires & Sea Dragons* is an accessible read as well as welcome leavening to bookshelves over-stuffed with popular dinosaur books. But if despite McGowan's book you are ever tempted to accommodate the idea of dancing dinosaurs at face value, it is worth bringing Disney's dancing hippopotami to mind. □

Henry Gee is on the editorial staff of *Nature*.

Molecular messengers

Jamshed R. Tata

Hormones. Edited by Etienne-Emile Baulieu and Paul A. Kelly. *Hermann*, 293 rue Lecourbe, 75015 Paris, France: 1990 Pp. 697. FF660.

It is no longer easy to write a comprehensive text on hormones. Endocrinology has ceased to be a self-contained discipline for physiologists and pathologists; it is a facet of regulatory biology of increasing interest to cell and molecular biologists, oncologists and also biotechnologists. In their preface, Baulieu and Kelly, the editors of *Hormones*, state that endocrinology has entered a new era and therefore "students, researchers, and physicians need a succinct source of classical knowledge coupled with recent information, findings and concepts". To what extent have they achieved this laudable goal?

In some ways, this is an unusual multi-author work. The first chapter, written by Baulieu and entitled 'Hormones: A Complex Communication Network', is a veritable monograph in itself. Its 164 pages, 187 figures, 34 tables and 212 references present an admirable synthesis of what is known about the formation, metabolism, action and clinical aspects of well-known vertebrate hormones. It appropriately focuses on hormone receptors, and the discussion of oncogenes places hormones in their proper perspective of intercellular signalling, and ends with a brief account of comparative and evolutionary aspects of endocrinology. The latter ironically draws attention to the lack of coverage of plant and invertebrate hormones in the rest of the book. There is a welcome attempt in the other 13 chapters to break away from the tiresome tendency of traditional endocrinology texts which divides topics according to organs in which different hormones are made. Instead, chapters deal with hormones according to the site of their production (adrenal medulla, thyroid and pancreatic hormones), or similarity of chemical structures (steroids, glycoprotein hormones) or common modes of action (calcium regulatory hormones, steroids).

An innovation is the inclusion of 17 short essays by distinguished scientists (five Nobel

laureates among them), interspersed between individual chapters. These turn out to be a mixed bag. Some convey the impression of 'name-dropping', some partially serve to fill the major gaps in the book (melatonin, prostaglandins), whereas others are indeed little gems. Among the latter are those by J. Glowinski on chemical messages of nerve cells, D. Gospodarowicz on growth factors (which is most apposite next to the chapter on growth hormone) and L. Birnbaumer on adenylate cyclase and G proteins, all packed with an impressive amount of information in only five pages. What these essays do collectively is to show convincingly that endocrinology can no longer be a self-contained discipline and that hormones are part of a wide array of signalling molecules highly conserved through evolution as key components of complex intercellular communication networks.

After Baulieu's comprehensive survey, the 13 chapters that follow are heterogeneous in their quality and coverage — not an uncommon feature of multi-author works. The material in the second chapter of only six pages on the anatomy of the hypothalamus—pituitary complex could have been incorporated in the next four chapters on pituitary and hypothalamic hormones, which comprise 118 pages. All four give a balanced account, without unnecessary detail, of the structure, synthesis, secretion, the possible mechanisms of their action and human disorders related to these hormones. I particularly liked Kelly's account of growth hormone and prolactin in one of these chapters. E. Milgrom's survey of steroid hormones is a veritable *tour de force*, being more informative and readable than the recent spate of specialist monographs on different aspects of these hormones. The chapter on clinical endocrinology of reproduction by S. S. C. Chen nicely complements Milgrom's, both authors stressing the importance of receptors in the molecular action and clinical disorders of reproductive hormones. Less emphasis is laid on receptors in the next chapter on the pancreatic hormones, insulin and glucagon. This is a pity, because insulin-resistance syndrome in man was among the first receptor mutation-based diseases in the rapidly growing list of hormone-resistance syndromes. Gastrointestinal hormones, angiotensin, aldosterone and atrial natriuretic factor, and the calcium regulatory hormones, are all comprehensively, yet succinctly, covered in the last three chapters.

How does Baulieu and Kelly's *Hormones* compare with other works on the same subject? Most recent multi-author publications deal with narrower, specialized topics concerning hormones, such as receptors, neuro-hormones and bone metabolism, so that it is futile and unfair to make an objective comparison with these. I shall therefore restrict myself to two other books with the same title. The first, in French, and also edited by Baulieu, can be considered as some sort of 1978 forerunner of the present book. Many

of the francophone authors of the 1990 *Hormones* had also written on the same or similar topics in the earlier book, with Baulieu also contributing the first overview chapter. But this new book is not simply an English version of the earlier French publication. Besides the 17 short essays, this book contains at least two to three times more material, with new chapters or expanded discussion on topics such as adrenocorticotrophic hormone, atrial natriuretic factor and clinical endocrinology of reproduction. The other *Hormones*, by A. Norman and G. Litwack (Academic, 1987: for review see *Nature* 329, 774; 1987), is not a multi-author production and is therefore more coherent in style. It was also restricted to vertebrate endocrinology but gave a more comprehensive coverage of prostaglandins and hormones of the pineal gland, kidney and thymus. Whereas Norman and Litwack had aimed their book primarily at medical students, Baulieu and Kelly's *Hormones* is more useful as a reference work for a wider readership of graduate students, basic and clinical investigators. It is also better presented and easier to use than both previous *Hormones*.

Perhaps the single major criticism of this book is that parts of it suffer from being out of date. There are relatively few references to post-1986 literature in many chapters. Although this would not significantly affect such topics as hormone structure, synthesis, secretion, metabolism or the biochemical basis of endocrine disorders, it is a drawback when seeking information about current views on mechanisms of hormone action, particularly regarding the nature and function of receptors and second messengers. So much is happening so rapidly in this area. If in the future one has to have a single book dealing with hormones, or if there is to be a second edition of *Hormones*, I would urge the authors not to restrict it to vertebrate hormones but to give adequate coverage to plant and invertebrate hormones and the developmental actions of hormones.

Hormones is well presented and, although it is a multi-author book, there is a set pattern of dividing the material within individual chapters. Its illustrations, over 600 in two colours, are particularly clear, informative and relevant to the text. For the clinical investigator there is an adequate discussion of both the traditional and the more trendy molecular basis of endocrine function and diseases in man. The references to literature are numbered and appear on the page on which they are first mentioned in each chapter rather than at the end. It also has a useful subject index. Baulieu and Kelly's *Hormones* is particularly recommended to those who wish to discover in a single volume a large diversity of facts concerning the nature, formation, function and relevance to man of the major vertebrate hormones. □

Jamshed R. Tata is at the National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

Environmental effects from burning oil wells in Kuwait

K. A. Browning, R. J. Allam, S. P. Ballard, R. T. H. Barnes, D. A. Bennetts, R. H. Maryon, P. J. Mason, D. McKenna, J. F. B. Mitchell, C. A. Senior, A. Slingo & F. B. Smith

Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ, UK

Model calculations, constrained by satellite observations, indicate that most of the smoke from the oil fires in Kuwait will remain in the lowest few kilometres of the troposphere. Beneath the plume there is a severe reduction in daylight, and a daytime temperature drop of $\sim 10^\circ\text{C}$ within ~ 200 km of the source. Episodic events of acid rain and photochemical smog will occur within $\sim 1,000$ – $2,000$ km of Kuwait. But changes in the Asian summer monsoon are unlikely to exceed the natural interannual variability and stratospheric ozone concentrations are unlikely to be affected.

THE main environmental effects arising from the burning oil wells in Kuwait come from emissions of oxides of carbon, sulphur and nitrogen, and unburnt hydrocarbons and particulates. The scenario used for this study¹, produced after discussion with the UK Department of Energy before the oil wells had been set alight, assumes 80 Tg of oil burning over a year, equivalent to the annual pre-invasion production of about 1.5 million barrels per day, with 6% of the oil being converted to fine particulate smoke. Small² assumed 1.6 million barrels per day, with 8.6% of the oil being converted to smoke. The products resulting from our scenario are given in Table 1, together with some current national and global emission figures for comparison.

We have used a simple model of plume rise to investigate the initial behaviour of the plume. We then used the Meteorological Office's mesoscale weather-prediction model to estimate the additional increase in height resulting from solar heating, the plume shape, and the daytime reduction in surface temperature, within a few hundred kilometres of Kuwait. For greater distances, out to several thousands of kilometres, we modelled the plume with the Meteorological Office's long-range dispersion model, which also provides estimates of wet and dry deposition. We also used data from a trajectory model as a basis for general circulation model integrations to determine the effect of the plume on the Asian summer monsoon.

Our study indicates that the smoke plume is very unlikely to reach the stratosphere, and the bulk of the smoke will remain within the lower troposphere before being deposited on the ground within a week or so of its emission. Consequently, changes in the Asian summer monsoon resulting from the smoke are likely to be small compared with the natural interannual variability. Beneath the plume within about 200 km of the source, it is likely that there will be a reduction in light to near night-time levels, and a reduction of the daytime maximum temperature of $\sim 10^\circ\text{C}$. The plume may remain visible for up to a few thousand kilometres, and slight reductions in daylight and temperature can be expected beneath it at distances of several hundred kilometres. Episodes of severe acid rain and 'black' snow, and photochemical smog, comparable with the highest

concentrations encountered around major industrial areas, are likely to occur out to distances of a thousand kilometres or more downwind. The carbon dioxide release (1% of current global emissions) will have an insignificant effect on global warming.

In general, we agree with Small² that effects will be significant in the Gulf region but insignificant on a global scale.

Local effects

During the first few minutes after emission, material in a smoke plume will rise rapidly, driven mainly by the heat from the burning oil. If we assume that the initial plume rise from each well is not enhanced by the effect of other wells, and use a heat source based on that expected from an average oil well (up to 500 MW), then simple calculations for turbulent plumes³ indicate that the smoke will reach an initial height of between 1 and 2 km in the range of stable conditions typical of winter. (Small² suggested a height of 700 m, but that was using a heat source of 319 MW.) In summer, when the well-mixed surface boundary layer deepens, smoke can be expected to rise to the top of the boundary layer, which can be 3 km deep (occasionally 4 to 5 km). In unstable situations, that is, in the presence of thunderstorms, the smoke could rise significantly higher, and some smoke might reach the upper troposphere; however, analysis of the Meteorological Office's climatological archives shows that these situations occur infrequently and are generally of short duration.

After the initial rise of the plumes they spread out and merge together. A further slow rise, termed 'self-lofting'⁴, may occur through solar heating. The consequence of this effect on the overall plume was investigated by using the Meteorological Office's mesoscale numerical weather-prediction model⁵, which has a gridscale of 15 km, adapted to include transport and diffusion of smoke in a manner similar to that developed for nuclear winter studies⁶. Transport of smoke by sub-gridscale convection, and removal by wet and dry deposition, were not included. In this and the other numerical models described later, incident solar radiation was assumed to be absorbed exponentially by the smoke with an absorption coefficient^{7,8} of $10\text{ m}^2\text{ g}^{-1}$. (Small² assumed $6.5\text{ m}^2\text{ g}^{-1}$ with a slightly greater amount of smoke.) The impacts of smoke on both scattering and infrared radiation were ignored. Nuclear winter studies⁷ show that neglecting the infrared properties of the smoke is a reasonable approximation when the particle sizes are predominantly less than one micrometre. The small infrared effects will act to ameliorate the predicted daytime surface temperature reduction and may reduce nocturnal surface cooling. The reduction of solar radiation by overlying cloud was also neglected, except at the surface, to simplify the model and to investigate the maximum possible impact.

Initial and boundary conditions for the mesoscale model were interpolated from the UK Meteorological Office's operational regional model forecasts. Throughout the period of the forecast, smoke was injected uniformly from the surface to a specified plume top. Forecasts were run for five selected days in January, February and March 1991 with and (as a control) without the interactive effects of absorption of solar radiation by smoke. We

selected cases to cover the two main synoptic situations: cloudy cyclonic, when the smoke was advected to the north over Iraq and Iran, and clear anticyclonic, when the smoke was advected to the southeast over Saudi Arabia and the Gulf. For January and February, the smoke sources were distributed uniformly over a square of side 45 km, from the surface to a height of 4 km. For the later cases, point sources of smoke were inserted from the surface to 2.5 km, at locations identified from satellite images. The plume height was deduced by comparing the forecast vertical wind profile with the direction of transport of the bulk of the observed smoke.

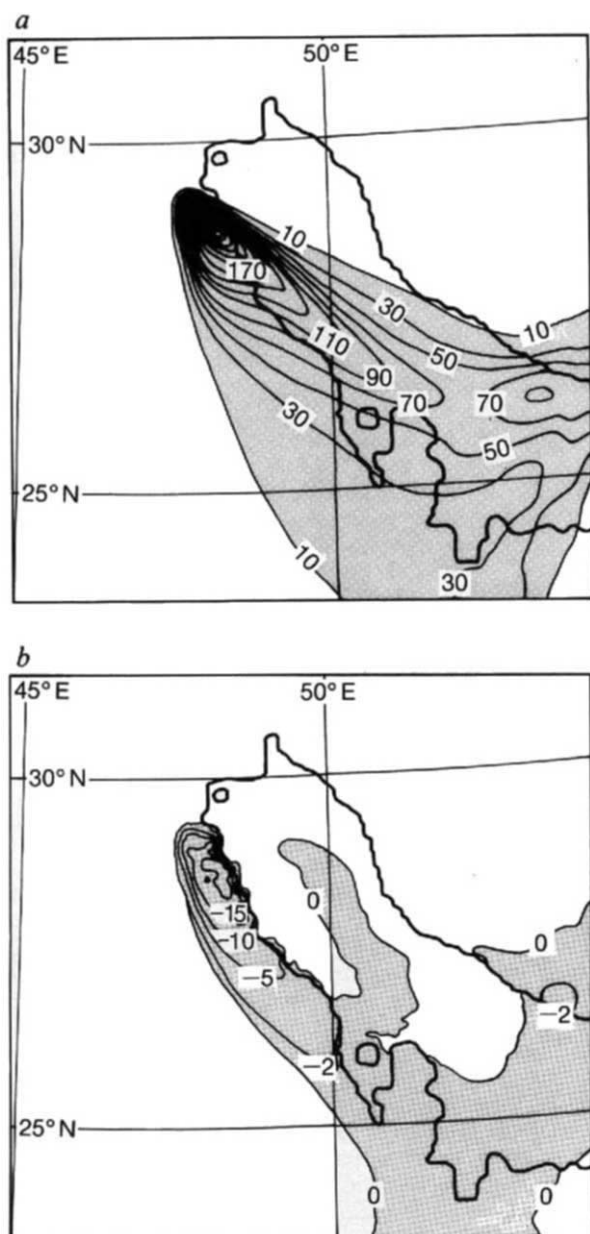


FIG. 1 Thirty-hour forecast of the Kuwaiti smoke plume valid at 12 GMT, 23 February 1991 (see also Fig. 2c), from the mesoscale model. *a*, Column-integrated distribution of smoke, S , in mg m^{-2} (optical depth $= S \times k \times 10^{-3}$, where the absorption coefficient $k = 10 \text{ m}^2 \text{ g}^{-1}$). *b*, Differences in screen-level (1.25 m) temperature due to absorption of solar radiation (contours at intervals of -5°C with contour added at -2°C). The smoke source, centred at 29.1°N , 47.8°E , is 5 megatons per year spread over a square of side 45 km, with its top at 4 km. The model domain is $1,350$ by $1,320$ km in the horizontal by 14 km in the vertical with a horizontal grid resolution of 15 km and 32 levels spaced irregularly in the vertical.

TABLE 1 Assumed annual production of the Kuwaiti fires

Type of emission	Amount (Tg per yr)	Comparison with current emission
Fine particulate black smoke*	5	Roughly a third of the carbon particles produced through tropical biomass burning ²²
Sulphur (as sulphur oxides)	2	Slightly more than the current UK annual sulphur emissions ²³
Nitrogen (as nitrogen oxides)	0.5	1988 UK nitrogen oxide emissions ²³ were 0.75 Tg
Carbon (ultimately as carbon dioxide)	60	About 1% of current global annual carbon dioxide emissions from fossil fuel combustion

* Small² assumes 5.8 Tg of smoke per year.

† Small (personal communication) suggests 3.3%.

The figures are compared with current national and global figures. The percentage of sulphur in Kuwaiti oil varies from field to field, but an average value is 2.5%† (figure provided by the Department of Energy). The percentage of nitrogen in the fuel is $\sim 0.17\%$ (Department of Energy), but to this must be added the quantity of nitrogen oxides produced during combustion (above $\sim 2,000$ K); we have assumed an overall figure of 0.6%.

In the control runs, the smoke remained below the top of the initial source for four of the cases. In the fifth forecast, south-westerly flow over the Iranian mountains lifted significant quantities of smoke from 2.5 to 4 km, and trace amounts reached 8–12 km in regions of convection. Peak concentrations of smoke were forecast just above the top of the boundary layer in all forecasts irrespective of the initial height of the source.

When the effects of solar absorption were included, lofting of smoke in significant quantities was limited to a rise of 1–3 km above that in the control integrations. With flow down the Gulf, and solar absorption included, small amounts of smoke reached 8 km. It is not clear whether this was due solely to self-lofting or partly due to the effects of induced mesoscale circulations, possibly through changing land-sea boundary effects (the model, in agreement with satellite imagery, sometimes generated cloud in the regions where smoke was elevated).

The predicted rise of the smoke is low compared with the typical height of the tropopause (15 km) over Kuwait. Therefore, with the smoke failing to reach the stratosphere, its lifetime can be expected to be relatively short (see later), and close analogies with nuclear winter^{9–11} and large volcanic eruptions¹¹ are inappropriate.

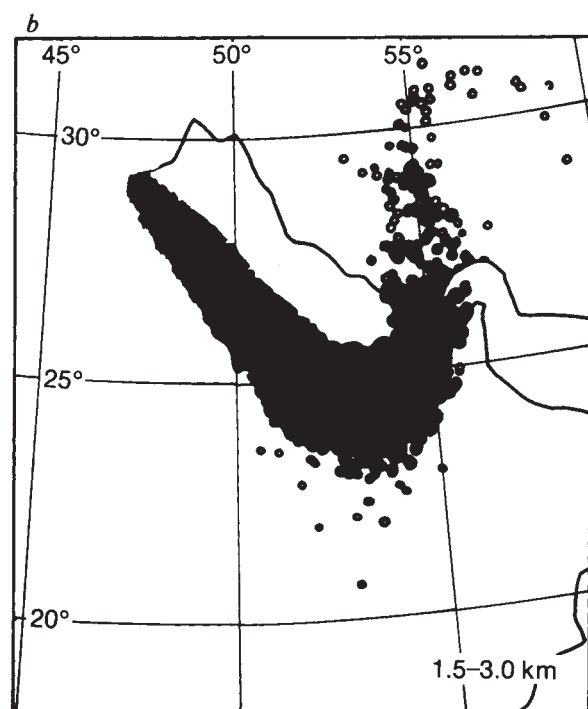
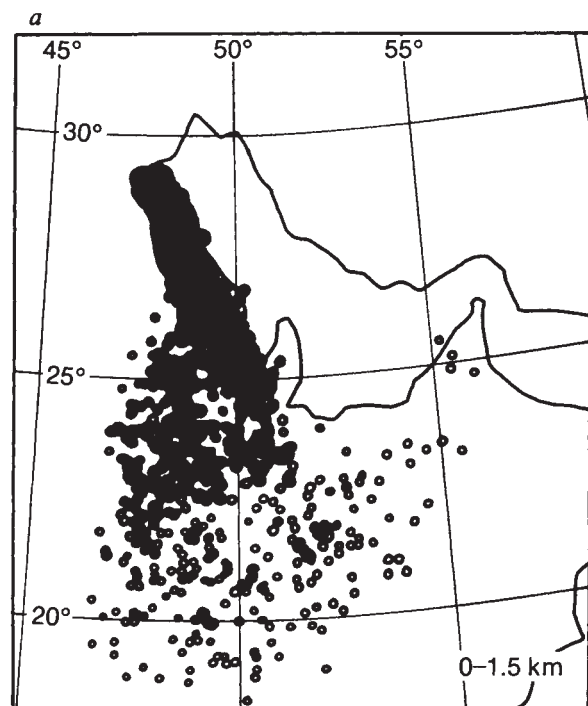
We also assessed the impact on surface temperatures of absorption of solar radiation. There is a large variability in the position of the plume, and the associated cooling and reduction in light levels, because of the day-to-day variability of the wind. Figure 1a shows the smoke distribution (compare also with Fig. 2 later), and Fig. 1b gives the associated reduction of surface temperature for 23 February 1991. The maximum daytime temperatures at screen level are depressed by about 10°C beneath the high concentrations of smoke within about 200 km of the source, and temperature drops of one or two degrees are predicted hundreds of kilometres downstream from the source. The drop of 20°C shown in Fig. 1b near the source region is likely to be excessive. The model cannot resolve the details of the flow on that scale, and the smoke concentrations may be large enough that infrared effects can no longer be ignored. Results will also be sensitive to the assumed surface characteristics and the modelling of the surface layer mixing. Sunlight is reduced to near night-time levels in the model when thick smoke is overhead, but the model cannot represent sub-grid-scale structure, and this may allow a little sunlight to penetrate.

From mid-March to mid-April, we have compared daily operational forecasts of midday temperatures in the absence of

the smoke with such observations as were available from stations beneath the plume (typically 4–8 observations per day when the plume extended south, very few when it extended north). Checks on the forecast values were provided by stations in unaffected regions. This subjective assessment suggests that within ~250 km of the source the reductions in midday temperature under the plume ranged from 5 to 8 °C, decreasing to 1–2 °C (probably the limit of accuracy) beyond about 750 km.

Effects at longer range

The Meteorological Office's long-range dispersion model¹² was used to investigate the longer-term distribution of the smoke, and also the acid rain and photochemical smog that may occur before concentrations reach more normal background levels.



The model simulates the spread of neutrally buoyant airborne pollutants (for example fine soot or sulphur) by releasing very large numbers of 'particles' into the model's atmosphere at the source of the pollution. The particles are transported by the model winds and vertical motions, and are diffused vertically and horizontally by random perturbations, which represent the small-scale turbulent motions likely to be experienced in the real atmosphere (especially near the surface). The model has been run daily since February to monitor the concentration of the smoke particles for a number of levels in the vertical. These integrations neglected self-lofting.

Six integrations were re-run with solar absorption included, as in the mesoscale model, and the particles allowed to self-loft in parcels to the level of neutral buoyancy. Comparisons between the integrations with and without solar absorption showed that in strong sunlight a small proportion (typically 5%) of the model's particles lofted rapidly through a kilometre or more. Thinning of the plume by entrainment, coupled with the model's normal turbulent diffusion, reduced the lofting to the point at which other vertical motions dominated, usually within one or two days. For an injection height of 2 km, the amount of smoke reaching 3–5 km was slightly augmented, but most of the material (90%) remained at lower levels.

Close to the source, results were similar to the mesoscale model. An example of the shape of the smoke plume at different heights is shown in Fig. 2a and b. In the boundary layer, below 1.5 km, the plume extends southwards and fans out; between 1.5 and 3 km, it encounters winds with a westerly component, and it drifts eastwards across the Gulf. Pictures from the geostationary satellite Meteosat, monitored with a dedicated computer display system, broadly confirmed the model forecasts. The visible image for the time corresponding to Fig. 2a and b shows the plume rapidly broadening towards the south as it

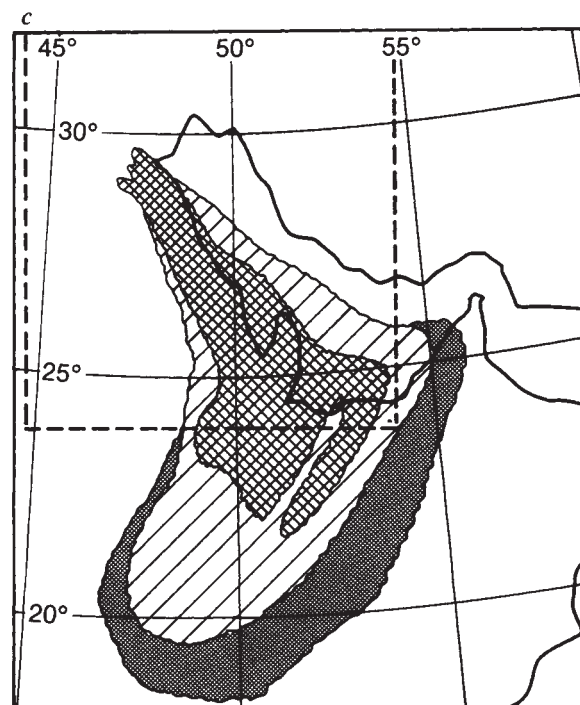


FIG. 2 Forty-eight-hour forecast, valid at 12 GMT, 23 February 1991, from the long-range dispersion model. a, Shape of the Kuwaiti smoke plume over the height interval 0 to 1.5 km; b, 1.5 to 3 km, as derived from the long-range dispersion model assuming that the smoke extended from 0.5 to 4 km altitude at source. These may be compared with the analysis of the Meteosat visible image for the corresponding time (Fig. 2c), where cross-hatched, hatched and stippled shading represent subjectively determined regions of different optical thicknesses. The dashed frame in c shows the area in Fig. 1.

encounters flows at different levels, before thinning to such an extent as to be undetectable (Fig. 2c).

To estimate the lifetime of the smoke, the 'mass' (corresponding to a small proportion of the source emission) carried by each model particle was progressively reduced to take account of the estimated wet and dry deposition to the surface. A standard deposition velocity for dry particles (0.05 cm s^{-1} , valid over the boundary-layer depth) and scavenging coefficient in rain ($1.2 \times 10^{-4} \text{ s}^{-1} \text{ per mm h}^{-1}$) were assumed for the smoke^{13,14}. Initial and boundary conditions, and rainfall rates, were taken from the operational analyses and forecasts of the Meteorological Office's regional model. Fourteen cases were considered. The half-life of the material was typically 5–15 days, except in very wet conditions, when it was shorter, and in deep dry boundary layers, when it was longer. Small² estimated a half-life of 5 days and MacCracken¹⁰ computed 3.5–6 days, depending on the depth of the release.

Although dry deposition is the principal process for removing gaseous sulphur dioxide from the plume, it is less important for 'sulphur' aerosol or sulphur attached to particulates, and the time-integrated dry deposition at a given location is unlikely to cause significant ecological damage, except possibly within the immediate area of the fires where the droplets of unburnt oil fall out under gravity. Wet deposition, although contributing less to the overall depletion of sulphur, will often be associated with locally very acidic rain. Measurements of pH in eastern Saudi Arabia¹⁵ taken before the Gulf War had a mean of almost 5.5, and therefore the associated background ions are unlikely to reduce the acidity generated by the plume significantly. We made calculations assuming plume widths (typically $\sim 200 \text{ km}$) and depths (3–4 km) based on our model results and satellite observations. Allowing for saturation in the uptake of acidic species into growing raindrops, the results indicate that, once rain is encountered, its acidity will depend only weakly on downwind distance out to at least 2,000 km, with pH values in the range 3.0–3.6 on many occasions. These very acidic values could damage crops and vegetation¹⁶.

Over the higher mountains much of the precipitation falls as snow and, where it is affected by the emissions, the snow may be blackened enough by soot to darken it considerably. Forecasts from the mesoscale model (such as in Fig. 1a) occasionally show average smoke-column densities of $\sim 40 \text{ mg m}^{-2}$ over the southern Iranian mountains. If this were scavenged by snowfall with a water equivalent depth of 1 mm, the mass mixing ratio

of soot within the snow would be typically 10^{-5} , which could double the absorption of solar radiation⁸. Spring snow-melt could therefore occur earlier and faster, with the possible consequence of high run-off and highly acidic pulses in the rivers.

Close to the source of the plume, the low light levels are expected to prevent any significant photochemical effects, except at the top of the plume. As the plume disperses, however, light levels will recover to values more typical of the region. This will allow rapid oxidation of hydrocarbons and other gases (such as H_2S) within the plume, probably over a timescale of a few days, resulting in locally severe episodes of photochemical smog.

Effect on the monsoon

To determine whether the smoke emissions are likely to affect the Asian monsoon, we carried out an idealized sensitivity experiment with a version of the Meteorological Office's atmospheric general circulation model that has been used extensively for seasonal predictions of tropical rainfall¹⁷; this version was chosen because it produces a good simulation of the Asian monsoon. The radiative effects of the smoke were included as in nuclear winter studies¹⁸; this approach was similar to the simple formulation in the mesoscale model described above, except that we included the effect of shading by clouds on the radiative heating of the smoke.

The distribution of smoke used in this model was derived from the wind fields actually observed in two previous years, 1989 and 1990, using the global trajectory facility attached to the numerical weather-prediction model of the European Centre for Medium-Range Weather Forecasting¹⁹. A smoke particle was released once a month from each of the vertices of an area of $1^\circ \times 1.5^\circ$ covering Kuwait at each of the four levels 950, 850, 700 and 500 mbar, and their three-dimensional trajectories were followed during an assumed 10-day lifetime. Figure 3a and b shows the resulting density of smoke at two different levels, averaged over the period of the monsoon, April to September.

For the sensitivity experiment, we prescribed purely absorbing smoke (absorption coefficient $10 \text{ m}^2 \text{ g}^{-1}$), with a smoke loading of 0.012 g m^{-2} (vertical optical depth 0.12) centred on Kuwait and mixed evenly throughout the troposphere up to a height of 8 km. The distribution of the smoke in Fig. 3 was approximated by $a^2/(a^2 + r^2)$, where r is the distance from source, and a was set at two model grid lengths ($\sim 600 \text{ km}$). The total smoke loading is equivalent to assuming a smoke lifetime of 10 days with emissions at a rate of 5 Tg per year.

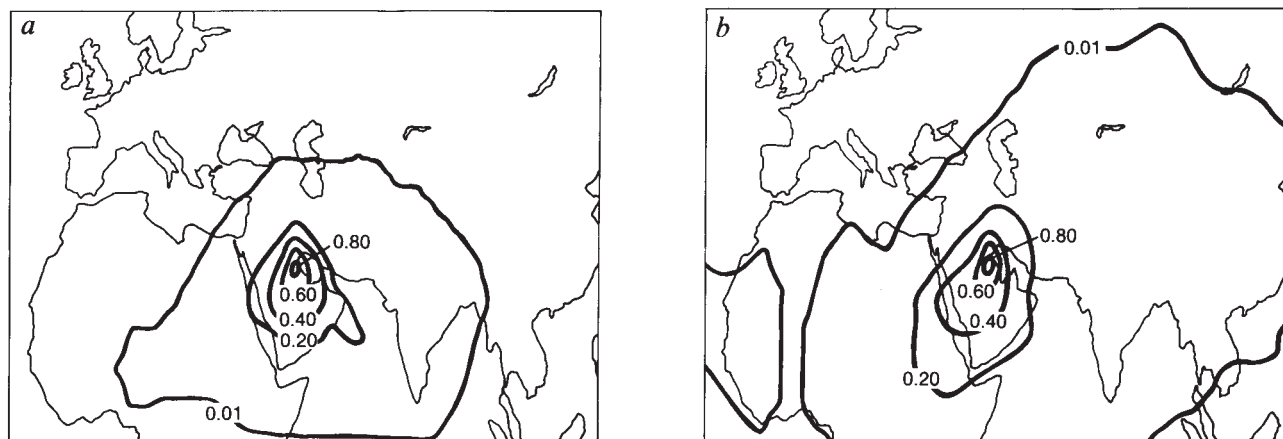


FIG. 3 Mean distribution of smoke for the period April–September derived from 10-day trajectories computed using the trajectory facility of the European Centre for Medium-Range Weather Forecasting¹⁹. The figure shows contours of the number of hours, during the 2-year period 1989–90, when trajectories, originating in Kuwait, covered various model grid squares, normalized by the corresponding number of hours over the Kuwait grid

square, for trajectories a, below 800 mbar and b, between 800 and 400 mbar. The contours may be roughly interpreted as normalized concentrations, although they exaggerate the amount of smoke at long range because they do not take into account losses by deposition. The gap and reappearance of material at longer range are due to the three-dimensional nature of the trajectories and the relatively small sample used.

The experiment and the control (with no smoke) were run from April to September. In the experiment, atmospheric solar heating in the centre of the plume increased by about 80 W m^{-2} (or 0.8°C per day), whereas the surface heating was reduced by $\sim 25 \text{ W m}^{-2}$. The surface and troposphere remained convectively coupled, so there was net heating of the troposphere-surface system which produced a slight enhancement of the Asian summer monsoon and, in most places, of the associated precipitation. This is in contrast to nuclear winter simulations in which the smoke is so thick and high that solar heating of the surface is effectively eliminated, the surface and troposphere are thermally decoupled²⁰, and there is a weakening of the Asian summer monsoon with reduced precipitation²¹.

Finally, we estimated the interannual variability of precipitation in the monsoon region, averaged over July to September, using data from existing simulations run with the observed sea surface temperatures for seven different years (1950, 1958, 1976, 1983, 1984, 1987 and 1988, encompassing wet and dry conditions in the Sahel¹⁷). The predicted changes in precipitation over southeast Asia ($5\text{--}30^\circ \text{N}$, $70\text{--}105^\circ \text{E}$) exceed twice the standard deviation of the interannual variations over only 10% of the area, and are all increases. In summary, the assumed scenario does not lead to significant decreases in the precipitation in the Asian summer monsoon, and might even lead to increases in certain regions. $\square$

Received 2 April; accepted 9 May 1991.

1. Note placed by Secretary of State for Defence in House of Commons Lib., 17 January 1991.
2. Small, R. D. *Nature* **350**, 11–12 (1991).
3. Briggs, G. A. *Plume rise* (US Atomic Energy Commission Rep. No. TID-25075 (US Department of Commerce, Springfield, Virginia, 1969)).
4. Radke, L. F., Lyons, J. H., Hobbs, P. V. & Weiss, R. E. *J. geophys. Res.* **95**, 14071–14076 (1990).
5. Golding, B. W. *Met. Mag.* **119**, 81–96 (1990).
6. Golding, B. W., Goldsmith, P., Machin, N. A. & Slingo, A. *Nature* **319**, 301–303 (1986).
7. Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B. & Sagan, C. *Science* **247**, 166–176 (1990).
8. Warren, S. G. & Wiscombe, W. J. *Nature* **313**, 467–469 (1985).
9. *Nature* **349**, 96 (1991).
10. *Science* **251**, 372 (1991).
11. Pearce, F. *New Scientist* **129**, 30–31 (1991).

12. Maryon, R. H. & Smith, F. B. *Österreich. Beiträge Met. Geoph.* **1**, 137–146 (1989).
13. Nicholson, K. W. *Atmos. Envir.* **22**, 2653–2666 (1988).
14. Jylhä, K. *Atmos. Envir.* **A25**, 263–270 (1991).
15. Ahmed, A. F. M., Singh, R. P. & Elmubarak, A. H. *Atmos. Envir.* **A24**, 2927–2934 (1990).
16. *UK Terrestrial Effects Review Group Rep. No. 1, Effects of acid deposition on terrestrial environments in the UK* (Department of the Environment, HMSO, 1987).
17. Folland, C. K., Owen, J., Ward, N. & Colman, A. J. *Forecasting* **10**, 21–56 (1991).
18. Mitchell, J. F. B. & Slingo, A. *J. geophys. Res.* **93**, 7037–7045 (1988).
19. Roskill, R. *ECMWF Met. Bull.* No. M1 1/6 (1987).
20. Cess, R. D. *Clim. Change* **7**, 237–251 (1985).
21. Ghan, S. J., MacCracken, M. C. & Walton, J. J. *J. geophys. Res.* **93**, 8315–8337 (1988).
22. Cachier, H., Burt-Menard, P., Fontugne, M. & Raucher, J. *J. Atmos. Chem.* **3**, 469–489 (1985).
23. *Acid deposition in the UK, 1986–1988*, Third Report UK Review Group on Acid Rain (Department of Transport, South Ruislip, UK, 1990).

Climate response to smoke from the burning oil wells in Kuwait

S. Bakan*, A. Chlond*, U. Cubasch*, J. Feichter†, H. Graf*, H. Grassl†, K. Hasselmann*, I. Kirchner*, M. Latif*, E. Roeckner*, R. Sausen†, U. Schlese*, D. Schriever*, I. Schult*, U. Schumann‡, F. Sielmann† & W. Welke*

* Max-Planck-Institut für Meteorologie, Bundesstrasse 55, D-2000 Hamburg 13, Germany

† Meteorologisches Institut der Universität, Bundesstrasse 55, D-2000 Hamburg 13, Germany

‡ Deutsche Forschungsanstalt für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, D-8031 Oberpfaffenhofen, Germany

The response of the global climate system to smoke from burning oil wells in Kuwait is investigated in a series of numerical experiments using a coupled atmosphere-ocean general circulation model with an interactive soot transport model and extended radiation scheme. The results show a decrease in surface air temperature of $\sim 4^\circ \text{C}$ in the Gulf region. Outside this region the changes are small and statistically insignificant. No weakening of the Indian summer monsoon is observed.

THE question of whether the smoke from the burning oil wells in Kuwait presents a serious threat to global climate has been the subject of considerable scientific debate and widespread public concern. Open oil burning produces a large quantity of soot particles, which absorb sunlight. This leads to a heating of the atmosphere in the absorbing soot layers and a reduction of the solar heating of the Earth's surface. It has been speculated¹ that this modification of the Earth's radiation balance could change the global climate, weakening the Indian monsoon and adversely affecting other climatically sensitive regions.

The effects of soot on the Earth's radiation budget have been

much studied in nuclear winter models². The present problem differs in several important aspects from these case studies. In particular, the magnitude and location of the pollution and the height of injection into the atmosphere are reasonably well known and can be expected to have a smaller impact on climate³ than the more extreme nuclear winter scenarios. The climate effects of emissions other than soot are probably negligible. The increase in carbon dioxide emissions, for example, relative to the existing emission level due to global fossil fuel use, is at most a few percent (Kuwait's pre-war contribution to world oil production⁴ was 2.6%).

Our study is restricted to the climate impact of smoke from burning wells on the global scale. Regional climate changes cannot be studied in detail with the relatively coarse resolution ($\sim 5^\circ$) of the available global models. Other environmental questions, such as the influence on stratospheric ozone or other trace gases, and in particular the impact on the ecology, are also not considered.

The model

The model set-up is an extension of the Hamburg global coupled ocean-atmosphere general circulation model which has recently been used in a number of simulations of greenhouse warming^{5,6}. The atmospheric component⁷ is a low-resolution version of the forecast model of the European Centre for Medium-Range Weather Forecasting which has been extensively modified for climate applications (for example, through the inclusion of

§ To whom correspondence should be addressed.

liquid water as a prognostic variable⁸). The model uses a dual spectral/grid representation with a triangular truncation at wavenumber 21 (5.6° horizontal resolution), has 19 vertical levels and includes the diurnal cycle and standard physics, such as cloud-radiation interactions⁸. It has been extensively validated and applied in climate studies^{9–11}. The ocean component is the 11-layer Hamburg 'large-scale geostrophic' circulation model, with a horizontal resolution matched to the atmospheric model. This has also been extensively validated and applied in a number of climate experiments^{12–14} and provides the core for the Hamburg-Scripps model of the ocean carbon cycle^{15–17}.

To compensate for the unavoidable initial climate drift of the coupled model, a standard flux correction¹⁸ is applied. This is equivalent to coupling only the deviations from the (not completely matched) initial equilibrium states of the individual subsystems and has no impact on the computed response for small climate perturbations.

The soot distribution is computed with an additional interactive tracer module¹⁹. This includes the processes of advection by the instantaneous three-dimensional velocity field, diffusion, vertical exchange through convection, washout by rain and dry deposition at the surface (assuming a deposition velocity of 0.1 cm s^{-1}). The wet removal rate is a nonlinear function of the local precipitation rate²⁰, which was originally tuned to yield a lifetime of 10 days for highly soluble aerosols (^{210}Pb and ^7Be). As the solubility of soot produced by uncontrolled oil burning, which is strongly contaminated by other substances, is not well known, the relation was retuned, in keeping with the worst-case philosophy adopted in this study, to yield a residence time of 20 days. All meteorological fields occurring in the computations of soot transport and removal are determined at each point and time by the atmospheric model.

We calculated the feedback of the soot distribution on the model's radiation balance using an extended radiation code based on Mie scattering theory for a size-dependent distribution of (spherical) aerosol particles^{21,22}. The computations are limited to the solar radiation fluxes, as the changes in the thermal infrared radiation were found to be negligible. Although soot particles are not spherical, so that Mie scattering theory is not strictly applicable, the computed net absorption coefficient ($\sim 8 \text{ m}^2 \text{ g}^{-1}$) for the size distribution assumed in the standard scheme (a zero-order log normal distribution with a mode radius²³ of $0.02 \mu\text{m}$) is in order-of-magnitude agreement with measurements (ref. 2 and Fig. 1). The effect of soot particles on

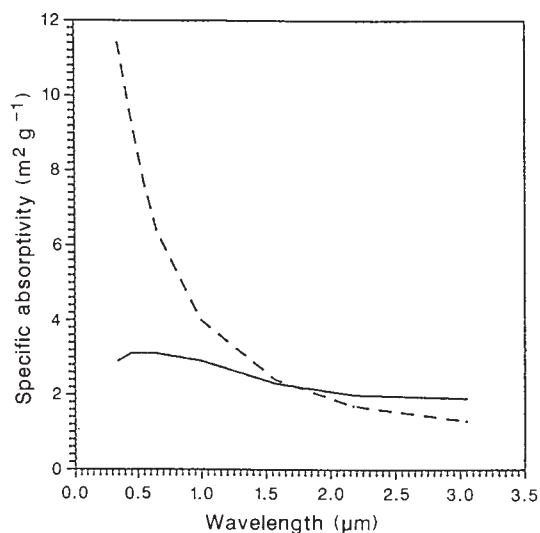


FIG. 1 Specific absorptivities of soot particles ($\text{m}^2 \text{ g}^{-1}$) as function of wavelength for two different radius scales for zero-order log-normal particle distributions. Dashed line, $r_m = 0.02 \mu\text{m}$; solid line, $r_m = 0.1 \mu\text{m}$. The value of $0.02 \mu\text{m}$ (radius at the peak of the distribution) was used in the standard scenario.

clouds, through the increase in cloud condensation nuclei or through the enhanced absorption of radiation by cloud droplets containing soot particles²⁴ could not be investigated with the present model.

The worst-case scenario

A number of climate simulations were carried out, but we shall present results only for the standard (worst-case) scenario. In this experiment, a constant burning rate of 440,000 tonnes oil per day (~ 3.1 million barrels per day) is assumed, corresponding to twice the pre-war production of Kuwait⁴. This is less than the considerably higher burning rates frequently quoted by Kuwait sources, but about 50% higher than calculations based on Kuwait oil-well data and estimates of reservoir pressure³. The burning is assumed to begin on 15 February and continue until the following January. The model was initialized by a 100-year integration (which had been made as a control run for other experiments addressing greenhouse warming⁵).

We assume that 10% of the oil (mass ratio) is converted to soot². This represents a reasonable mean value of experimental data for burning oil pools, but is probably rather high for better-ventilated burning oil wells. The soot is introduced into the atmosphere at a single model gridpoint, but some spatial smoothing is applied to avoid negative concentrations arising

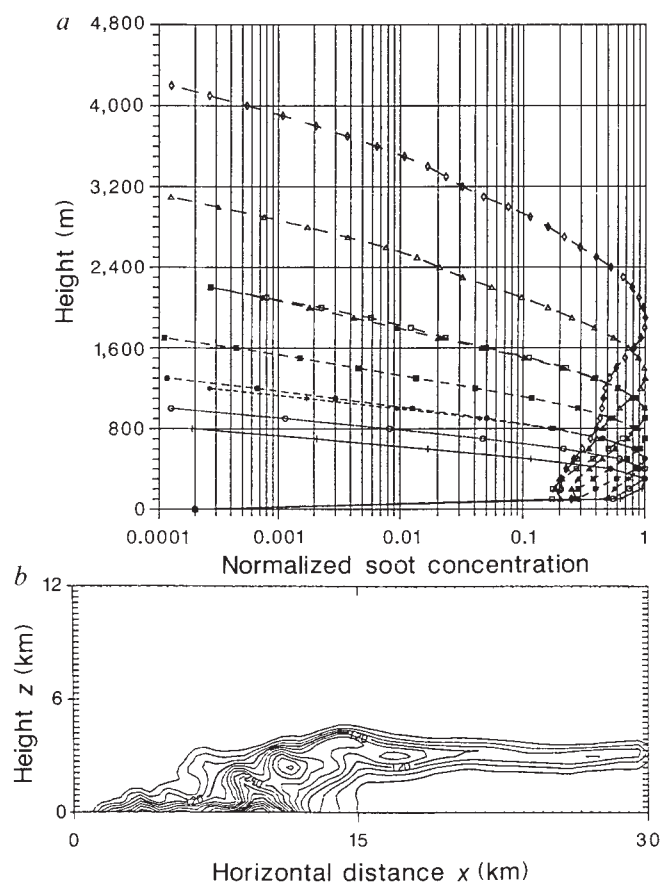


FIG. 2 Estimate of the injection height of the soot using mesoscale circulation models. *a*, Soot profiles for different energy densities, E , and different static stabilities, σ , in the simulation with the first model²³. Runs 1 ($+$), 2 ($\circ$), 3 ($\times$) and 4 ($\bullet$), $E = 5 \times 10^8 \text{ J s}^{-1}$; runs 5 ($\square$) and 6 ($\triangle$), $E = 2 \times 10^9 \text{ J s}^{-1}$; runs 7 ($\square$), 8 ($\triangle$) and 9 ($\diamond$), $E = 8 \times 10^9 \text{ J s}^{-1}$. Runs 1, 4 and 7, $\sigma = 2.7 \times 10^{-3} \text{ K m}^{-1}$; runs 2, 5 and 8, $\sigma = 1.3 \times 10^{-3} \text{ K m}^{-1}$; runs 3, 6 and 9, $\sigma = 6.7 \times 10^{-4} \text{ K m}^{-1}$. The energy was released in an area of 4 km^2 . The profiles are normalized by the maximum concentrations. *b*, Height section of the soot concentration in the simulation with the second model²⁴. The contour interval is 0.3 mg per kg .

from the numerical truncation of the spectral model. To allow for the rapid sedimentation of large soot particles within the first 500-km grid cell, the source strength is reduced by 15%, so that the net effective soot injection rate on the model scale is 37,400 tonnes per day.

The soot is injected initially at a constant mixing ratio into the lowest two kilometres of the atmosphere (more precisely, into model levels 2–5, between 60 m and 1,800 m). The initial injection profile is based on mesoscale simulations of 'nuclear winter'²⁵, on our own computations using two different three-dimensional, nonhydrostatic mesoscale models and on satellite observations. If all the oil is assumed to be burned in an area of $20 \times 20 \text{ km}^2$, the energy release, assuming a typical energy equivalent for oil of $4 \times 10^4 \text{ J kg}^{-1}$, yields a heat flux of 500 W m^{-2} . According to the mesoscale model simulations for nuclear winter²⁵, this results in injection heights, depending on the atmospheric stability, of about 2 km. Figure 2a shows the corresponding injection heights computed with one of our mesoscale models²⁶ for various values of the atmospheric static stability and surface heat release. These curves, as well as the results of a second mesoscale model²⁷ (Fig. 2b shows an example of a soot plume computed with this model) indicate that for the range of static stabilities typical of the Gulf region, in which unstable deep cumulus convection events are rare, the soot injection will normally not exceed a height of 2–4 km, even for the relatively large heat fluxes assumed in our simulations. This is supported, finally, by an analysis²⁸ of a satellite observation of the soot cloud over Kuwait on 24 February 1991, which indicated that the cloud did not exceed a height of 3 km.

The relatively low injection height is an important factor in limiting the global climate impact of the soot release³. The soot is transported in atmospheric layers in which it can be rather rapidly washed out and removed by dry deposition. The 20-day residence time of soot in the troposphere is small compared with typical residence times of aerosol in the stratosphere, which are several years. Also, the opposing effects of atmospheric heating in the absorbing soot layers and cooling of the Earth's surface can be more easily compensated in the lower troposphere through vertical energy transfer.

The soot distribution becomes quasi-stationary one or two months after the oil burning begins on 15 February. The total atmospheric soot content in the quasi-stationary state is $\sim 7 \times 10^8 \text{ kg}$. The soot does not rise significantly higher than three kilometres. The highest altitudes are attained in the summer, when the increased solar radiation increases the convective activity. The self-lofting of the soot (the buoyancy uplift caused by the absorption of solar radiation in the layers of a high soot concentration) is also largest in the summer. But numerical experiments with a mesoscale model²⁶ indicate that this effect is small compared with the turbulent mixing. The transport of soot into the stratosphere is negligible: $\sim 0.3\%$ (2,100 tons) of the total atmospheric soot load is found in the stratosphere.

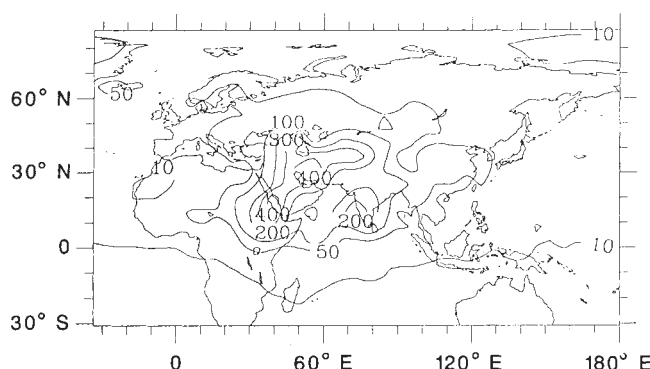


FIG. 3 Total deposition of soot (mg m^{-2}) during the 12-month simulation. Contour intervals are at 10, 50, 100, 200, 300 and 400 mg m^{-2} .

Most of the soot is deposited, primarily through wash-out by rain, within a few thousand kilometres from the source (Fig. 3). The maximum surface densities of $\sim 400 \text{ mg m}^{-2}$ after one year of continual burning are found in Iran (the assumed deposition of 15% of the soot emission in the immediate vicinity of the oil wells, which can yield densities two to three orders of magnitude larger than this value, is not shown in Fig. 3).

The climate response to the atmospheric soot loading was determined as the difference between the worst-case scenario and the control run. As expected from Fig. 3, considerable anomalies in the radiation field are found only on regional scales within a few thousand kilometres from the source. The reduction off the net solar radiation reaching the Earth's surface in this region is $\sim 20\%$, while the net absorbed solar radiation by the atmosphere and Earth's surface is increased by 15%. The total optical depth of the cloud, including absorption and scattering, is ~ 0.4 .

The surface deposition of soot also affects the surface albedo, in particular for snow-covered surfaces. Soot that is covered by fresh snowfall in the winter reappears during the melting period, when the accumulated soot concentration can increase to values of 10 p.p.m. This can reduce the albedo of the snow from 60% to $\sim 40\%$ (ref. 29). The area of polluted snow is small, however, and the duration of the melting period is short. On the basis of previous sensitivity experiments¹¹ with changed snow covers, we regarded the albedo effect as negligible and did not include it in the radiation computations.

It is well known from El Niño/Southern Oscillation (ENSO) studies¹⁰ and other atmospheric response experiments that regional-scale heat flux or radiation anomalies can generate global-scale atmospheric anomalies. But there is no obvious large-scale structure evident in either the temperature or the precipitation fields that is distinguishable from the inherent model interannual variability (r.m.s. monthly mean deviations of $\sim 2^\circ\text{C}$ in temperature and 50%, but very small scale, in precipitation^{5,8}). On the regional scale, significant temperature anomalies are seen mainly near the source region (Fig. 4, left), with a maximum cooling of $\sim 4^\circ\text{C}$ in July. All values shown represent averages over at least a model grid square of side 500 km and a period of one month. The temperature decrease can be much stronger in smaller regions and timescales below dense smoke plumes. Outside the Gulf region, considerable anomalous heating is found over the Tibetan Plateau in May. This is probably due to the proximity of the absorbing soot layers in the atmosphere to the surface at these high altitudes (after May, the soot trajectories move further south).

In the precipitation distributions (Fig. 4, right), there is no indication of a weakening of the Indian summer monsoon, as has frequently been speculated. There is even a region of enhanced rainfall over India during the onset phase of the monsoon in May, in response to the anomalous heating over Tibet. The enhanced rainfall region shifts further north during the following two months.

The conclusion that, on the global scale, the anomaly fields cannot be distinguished from the natural model variability is supported by a lack of agreement between the global response patterns found in the standard scheme and a number of further sensitivity runs made with modified input and model parameters.

Sensitivity studies

To investigate the dependence of the simulation results on some of the more critical input and model assumptions, we carried out a series of additional experiments. First, the value for soot emission was changed. Halving the emission roughly halved the amplitude of the climate response in the near-field regions, whereas the large-scale response was again indistinguishable from noise. But an alternative scheme, in which a number of input parameters were changed simultaneously to yield an effective increase of the emissions by an (unrealistic) factor of 4, produced a considerable climate response on the global scale,

with temperature decreases, and in some regions increases, of about 5–10 °C. Interpolating between the three simulations, we conclude that although the large-scale climate change in the worst-case scenario is smaller than the natural interannual variability, it is not far from the detectability threshold.

Because of the large uncertainty in the size and the resultant specific absorption of soot particles, we considered an alternative size distribution of soot particles. The particle distribution was rescaled to a mode radius of 0.1 μm instead of 0.02 μm , yielding a reduction of the specific absorption cross-section by a factor 2–4 in the relevant solar spectral band (Fig. 1). The near-field climate response decreased by a similar factor and was barely detectable above the natural variability. The true size distribution presumably lies somewhere between these cases², but because of the flakelike geometry of soot particles the first scenario is probably more realistic.

Further sensitivity experiments were carried out with modified injection heights and larger washout rates corresponding to residence times of about 10 days. The sensitivity with respect to these parameters was approximately linear. In summary, our standard model may be regarded as a reasonable worst-case

scenario within the range of uncertainties of the various input and model parameters.

Discussion

An advantage of the fully coupled model system used in this study is that all processes are computed interactively. Thus the effect on the soot transport of the thermal uplift induced by warming in the layers of maximal soot concentration is automatically included, as is the inhibition of the vertical transport of soot because of the higher static stability established between the heated layers and the cooled surface. Similarly, the complete feedback chain describing the influence of the changed circulation on the precipitation distribution and thus on the soot washout rate, and thereby, finally, on the radiation perturbations driving the circulation change, is incorporated into a single model.

The oceans, although also coupled fully interactively with the atmosphere, play a relatively small part in the present experiment. The deep ocean circulation does not change significantly in the short one-year integration period, and the changes in the more rapidly responding upper layers and in the sea ice distribu-

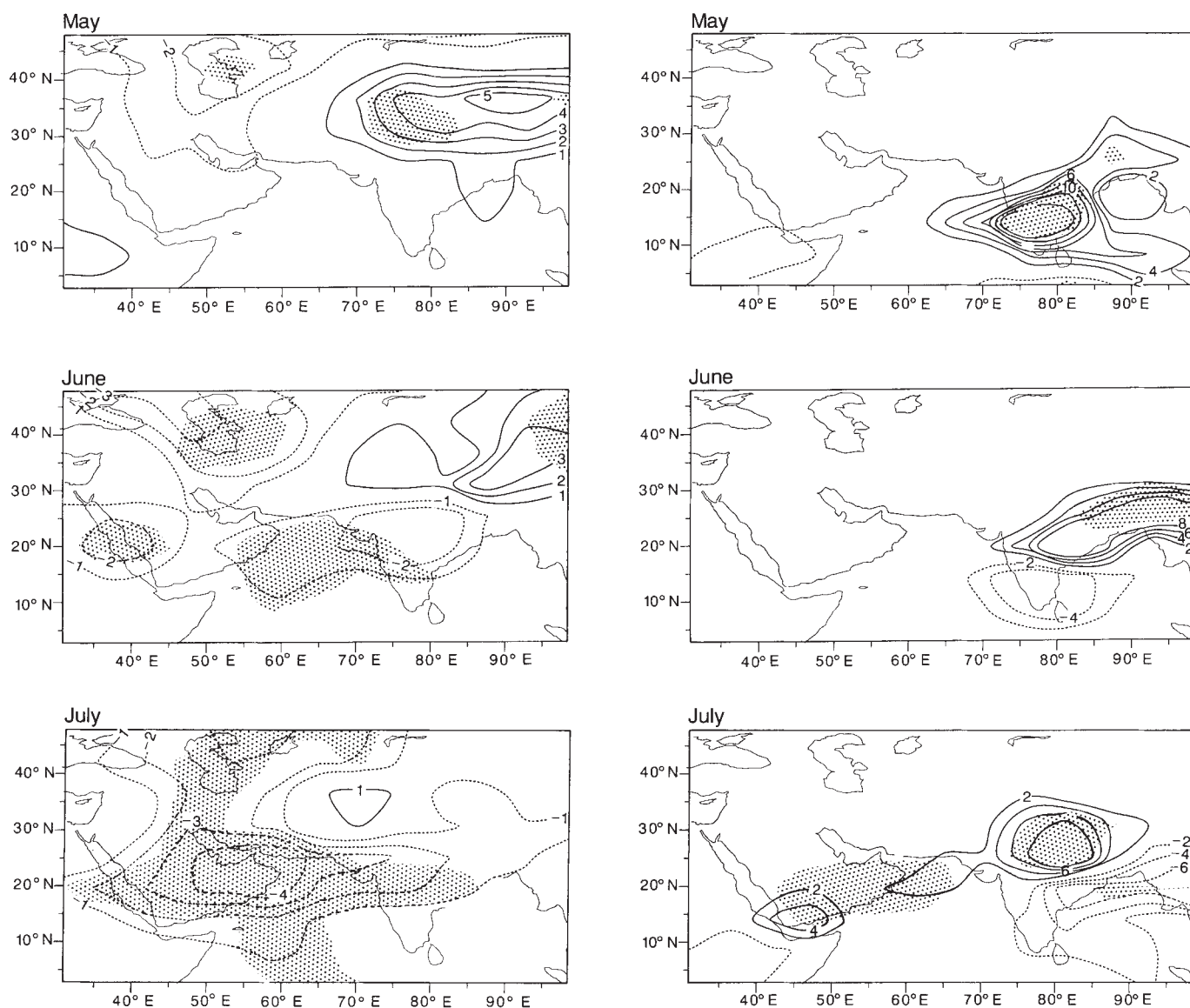


FIG. 4 Climate response in summer. Left: anomalous 2-m temperature. Contour interval is 1 °C. Right: Anomalous precipitation. Contour interval is

2 mm per day. The zero contours have been suppressed. Shading indicates regions with local significance levels larger than 95%.

tion occur in regions where the climate changes cannot be detected above the natural variability background.

The principal reason for the lack of a statistically significant global climate response outside the near-source region is the relatively low initial height for soot injection. This in turn is governed by the limited heat release and the finite static atmospheric stability. The low injection level enables the soot to be washed out or removed by dry deposition in a typical lifetime of 20 days. Only 0.3% is transported into the stratosphere, where soot can survive for several years.

The limited height of the soot distribution also enables the positive and negative anomalies in the radiation balance induced by the smoke to be more easily compensated through the turbulent energy exchange and infrared radiation fluxes in the lower troposphere: the increased downward longwave radiation from the layers of high soot concentration, which are warmed by the absorption of solar radiation, and the decreased upward turbulent heat fluxes from the surface, which is cooled by the reduced solar flux below the smoke cloud, counteract these

opposing temperature trends.

Although we were unable to detect a statistically significant climate signal on the global scale in our one-year standard scenario (or any of the other sensitivity experiments, apart from the unrealistic case in which the effective soot emissions were increased by a factor of four), this does not imply that such a signal cannot be found. By extending the experiment over a number of years, or carrying out an ensemble of such experiments, the persistent response pattern could presumably have been filtered out from the nonrecurrent natural interannual variability. But the fact that the signal could not be detected in a one-year integration implies that the global climate perturbation should have little impact in practice.

The large global climate response obtained when the effective soot emission was increased by a factor of four indicates, however, that our standard scenario could not have been very far from the detectability threshold, even for a one-year simulation, and may serve as a timely reminder that our global climate is indeed vulnerable to irresponsible actions by man. □

Received 1 April; accepted 8 May 1991.

1. Aldhous, P. *Nature* **349**, 96 (1991).
2. Pittcock, A. B. *et al. Scope* **28**, 1 (1986).
3. Small, R. D. *Nature* **350**, 11–12 (1991).
4. *Statistical Review of World Energy* (British Petroleum Company, London, 1990).
5. Cubasch, U., Sausen, R., Oberhuber, J., Lurkeit, F. & Böttinger, M. Simulation of the greenhouse effect with coupled ocean-atmosphere models *Cray Channels* **12**, 6–9 (1990).
6. Hasselmann, K. *Tellus* (in the press).
7. Dümenil, L. & Schlesse, U. Meteor. Inst. Univ. Hamburg. Rep. No. 1 (Hamburg, 1987).
8. Roeckner, E. *et al. WMO/TD No. 322* (WMO, Geneva, 1989).
9. Cess, R. D. *et al. Science* **245**, 513–516 (1989).
10. Latif, M., Biercamp, J., v. Storch, H., McPhaden, M. J. & Kirk, E. *J. Clim.* **3**, 509–521 (1990).
11. Barnett, T. P., Dümenil, L., Schlesse, U., Roeckner, E. & Latif, M. *J. Atmos. Sci.* **46**, 661–685 (1989).
12. U. Mikolajewicz, Santer, B. D. & Maier-Reimer, E. *Nature* **345**, 589–593 (1990).
13. U. Mikolajewicz & Maier-Reimer, E. *Clim. Dynam.* **4**, 145–156 (1990).
14. Maier-Reimer, E. & Mikolajewicz, U. *Experiments with an OGCM on the Cause of the Younger Dryas, Oceanography 1988* (eds Ayala-Castaneres, A., Wooster, W. & Yanez-Arancibia, A.) 87–100 (UNAM Press, Mexico City, 1989).
15. Maier-Reimer, E. & Hasselmann, K. *Clim. Dynam.* **2**, 63–90 (1987).
16. Heinze, C., Maier-Reimer, E. & Winn, K. *Palaeoceanography* (in the press).
17. Bacastow, R. & Maier-Reimer, E. *Clim. Dynam.* **4**, 95–125.
18. Sausen, R., Barthel, K. & Hasselmann, K. *Clim. Dynam.* **2**, 145–163 (1988).
19. Feichter, J., Roeckner, E., Schlesse, U. & Windelband, M. *Proc. 18th ITM on air pollution modelling and its applications* (ed. Van Dop, H.) (Plenum, New York, in the press).
20. Giorgio, F. & Chameides, W. L. *J. geophys. Res.* **91**, 14367–14376 (1986).
21. Bakan, S. thesis, Univ. Hamburg (1982).
22. Schult, I. thesis, Max-Planck-Inst. Meteorol., Hamburg (1991).
23. Espenscheid, W. F., Kerker, M. & Matijevic, E. *J. phys. Chem.* **68**, 3093–3097 (1964).
24. Grassl, H. *The Changing Atmosphere, Dahlem Workshop Reports* (eds Rowland, F. S. & Isaksen, I. J. A.) 187–199 (Wiley, Chichester, 1988).
25. Penner, J. E., Haselman, L. C. & Edwards, L. L. *J. Clim. appl. Met.* **25**, 1434–1444 (1986).
26. Chlond, A. *Boundary-Layer Meteorol.* (submitted).
27. Schumann, U., Hauf, T., Höller, H., Schmidt, H. & Volkert, H. *Beitr. Phys. Atmos.* **60**, 413–446 (1987).
28. Schumann, U. *DLR preliminary rep.* (DLR, Oberpfaffenhofen, 1991).
29. Warren, S. G. & Wiscombe, W. J. *J. Atmos. Sci.* **37**, 2734–2745 (1980).

ACKNOWLEDGEMENTS. This research was mainly supported by the Bundesministerium für Forschung und Technologie (BMFT). The computations were carried out at the Deutsches Klimarechenzentrum, which is supported by the BMFT, the Max-Planck-Gesellschaft, the Freie und Hansestadt Hamburg, the Alfred Wegener Institut für Polar- und Meeresforschung and the Forschungszentrum Geesthacht. We are grateful to M. Grunert for preparing the diagrams.

Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*

Titia K. Sixma, Sylvia E. Pronk, Kor H. Kalk, Ellen S. Wartna, Ben A. M. van Zanten, Bernard Witholt* & Wim G. J. Hol

BIOSON Research Institute and *Laboratory of Biochemistry, Department of Chemistry, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Examination of the structure of *Escherichia coli* heat-labile enterotoxin in the AB₅ complex at a resolution of 2.3 Å reveals that the doughnut-shaped B pentamer binds the enzymatic A subunit using a hairpin of the A2 fragment, through a highly charged central pore. Putative ganglioside GM₁-binding sites on the B subunits are more than 20 Å removed from the membrane-crossing A1 subunit. This ADP-ribosylating (A1) fragment of the toxin has structural homology with the catalytic region of exotoxin A and hence also to diphtheria toxin.

certain functional properties such as the ADP-ribosylation of GTPases. This family includes pertussis toxin, diphtheria toxin and exotoxin A^{1,2}.

The enterotoxins are important agents in diarrhoeal diseases of man (LT and CT) and animals (LT). LT and CT both consist of one A subunit with relative molecular mass ~27,000 (*M_r* ~ 27K), and five B subunits with *M_r* ~ 11.6K, forming an AB₅ complex of *M_r* ~ 86K in total (for reviews see refs 1–4). Polyclonal antisera raised against either LT or CT crossreact with the other toxin⁵ because the sequence of both the A and B subunits of these toxins is more than 80% identical (refs 6, 7, and reviewed in ref. 4).

Both the A and the B subunits of CT and LT are synthesized intracellularly as precursor proteins. After translocation across the bacterial cytoplasmic membrane and removal of leader peptides, the AB₅ complex is assembled in the periplasm^{8,9}. Some *Vib. cholerae* strains release toxin directly into the surrounding medium, whereas enterotoxigenic *E. coli* strains may release toxin as part of outer membrane fragments^{8,9}.

ENTEROTOXIGENIC *E. coli* and *Vibrio cholerae* strains produce 'heat-labile enterotoxin' (LT) and cholera toxin (CT), respectively. These proteins are part of a family of toxins that share

TABLE 1 Structure determination data

Dataset	Unique reflections	Complete (%)	R_{merge}^* (%)	$R_{\text{native}}^\dagger$ (%)	x	y	z	Occupancy	Anomalous occupancy	B-factor ($\text{\AA}^2$)	$R_{\text{cullis}}^\ddagger$ (%)	Phasing power
Native I	11,979	83.4	3.4	—								
Native II	29,769	85.8	6.2	6.2								
K_2PtCl_4	12,282	85.6	4.6	14.7	0.1855	0.0920	0.1560	1.01	1.04	aniso	69.8	1.45
					0.1077	0.3845	0.1391	0.40	0.39	31		
					0.1116	0.3522	0.8289	0.20	0.27	26		
					0.1327	0.6483	0.3109	0.23	0.25	41		
					0.1393	0.6118	0.7075	0.08	0.08	5		
KAuCl_4	13,177	91.8	3.0	14.5	0.1748	0.0396	0.1290	0.52	0.53	51	75.8	1.33
					0.2015	0.2141	0.7414	0.83	0.73	3		
					0.1866	0.0907	0.1583	0.42	0.40	1		
K_2OsO_4	13,382	91.8	3.6	14.3	0.1994	0.2153	0.7398	1.12	1.04	6	70.6	1.57
					0.1858	0.2036	0.9795	0.08	0.08	5		
					0.0178	0.9155	0.6744	0.05	0.02	5		

Heat-labile enterotoxin from porcine *E. coli* was overproduced, purified and crystallized as reported in ref. 55. The crystals have space group $P2_12_12_1$ with cell dimensions a , 119.2 Å; b , 98.2 Å; c , 64.8 Å. **Data collection:** all data used in the structure determination were collected on an Enraf Nonius FAST area detector, with an Elliot GX 21 rotating anode generator. All crystals except native II were oriented and rotated in exactly the same manner with the b^* axis along the ω axis to collect anomalous pairs optimally. Oscillation ranges were optimized for completeness with a data collection strategy protocol developed by I. Vickovic. Data processing was done off-line with MADNES (ref. 56) with the XDS profile fitting option⁵⁷. **Heavy atom derivatives:** crystals were soaked in a standard mother liquor (10% polyethylene glycol 6000, 175 mM KF, 100 mM Tris, 1 mM EDTA, 0.02% azide, pH 7.50) for 3 days for KAuCl_4 (3 mM in the dark), 18 h for K_2PtCl_4 (5 mM, but mounted in standard mother liquor) and 6 days for K_2OsO_4 (3 mM). A single heavy atom site could be easily identified from the K_2PtCl_4 difference Patterson map. Subsequent sites were found by cross-difference Fourier analyses. Heavy atom refinement with Phare (program provided by G. Bricogne) gave an overall figure of merit of 0.62 (15–3.1 Å). **Pentamer axis:** a solvent mask, calculated with the Leslie algorithm⁵⁸, showed a protein with a doughnut shape of dimensions similar to the electron microscopy data published for cholera toxin⁴² and a smaller (A) subunit. Fivefold features of the envelope were used for an initial estimate of the position of the fivefold axis. The four parameters defining this axis were then optimized by searching for a maximum in the density correlation coefficient using a spherical volume in the five B subunits. **Density modification:** a starting envelope for density averaging was created from the solvent mask by filling the protein volume region with random atoms. An averaging envelope was edited out with 'A subunit', 'B subunit' and 'solvent' regions. Fivefold averaging of the B subunits and solvent flattening were done simultaneously, whereas the A subunit was kept unchanged, using the program suite DEMON, written by F.M.D. Vellieux. The B subunit could be traced unambiguously in the averaged map. The protein structure was built on an Evans and Sutherland PS390 using the program FRODO (ref. 59) with help of skeletonized maps produced in 'Bones' (ref. 60). **The A subunit:** in alternating cycles of averaging and model building, with gradual phase extension from 3.1 to 2.7 Å, a model was produced of the A subunit with ~65% of the main chain traced correctly. The incomplete AB_5 model was then improved with molecular dynamics refinement using GROMOS (ref. 26). Phases from this model were combined with both MIR and averaged phases by the SIGMAA procedure²⁷, giving a map in which the A subunit could be traced. Further refinement was done by GROMOS and TNT (ref. 61), including restrained individual temperature factors. The current crystallographic R factor, defined as $\sum |F_{\text{calc}} - F_{\text{obs}}| / \sum F_{\text{obs}}$, is 17.8% between 8 and 2.3 Å, with r.m.s. deviations from ideality of 0.018 Å for bond lengths and 3.1 degrees for bond angles. In the Ramachandran plot six non-glycine residues are outside the allowed regions, five of which are in the poorly defined regions mentioned in the text.

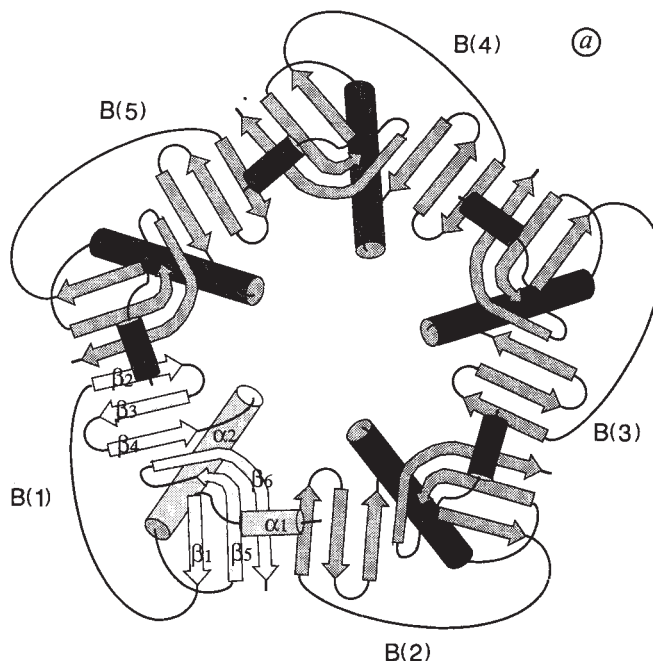
$$* R_{\text{merge}} = \frac{\sum_{hkl} \sum_{\text{refl}, j} |I(hkl, j) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{\text{refl}, j} \bar{I}(hkl)}$$

$$\dagger R_{\text{native}} = \frac{\sum |F_2 - F_1|}{\sum F_1} \text{ with } F_1 \text{ referring to native dataset I.}$$

$$\ddagger R_{\text{cullis}} = \frac{\sum ||F_{\text{PH}}| - |\bar{F}_{\text{P}} + \bar{F}_{\text{H(calc)}}||}{\sum ||F_{\text{PH}}| - |\bar{F}_{\text{P}}|} \text{ for centric reflections only.}$$

§ Phasing power is the mean value of the heavy atom structure amplitude divided by the residual lack of closure error.

After release from the bacterial cell, the toxin binds to intestinal epithelial cells of the host using the B subunits, which interact specifically with gangliosides G_{M1} on the outside of the membrane bilayer of the target cell (reviewed in refs 2, 3). Part of the A subunit is then transferred across the target-cell membrane. This requires that the A subunit be 'nicked' proteolytically between residues 192 and 195 (refs 7, 10). The single disulphide bridge between residues 187 and 199 must also be reduced^{11–13}. This results in an active A1 fragment of ~192 residues and an A2 fragment of ~45 residues. After translocation through the membrane, the catalytic site of the A1 fragment ADP-ribosylates an Arg residue of the G_{Sa} regulatory component of the adenylate cyclase system^{14–16}, a reaction which requires not only NAD as second substrate, but also GTP and another GTPase (also referred to as ADP-ribosylation factor or S protein^{2,17,18}). After ADP-ribosylation, the G_{Sa} subunit remains in a permanently activated state, which stimulates adenylate cyclase. This elevates levels of cyclic AMP which are responsible for the efflux of fluids and ions from the affected cell. The modified G_{Sa} protein may also exert effects through increased prostaglandin E levels¹⁹. Regardless of the details of the macromolecular cascade



induced, toxin action results in intestinal diseases ranging from mild traveller's diarrhoea to severe dehydration and eventual death, in particular in developing countries.

Many vaccines have been developed against cholera and other diarrhoeal diseases over the past two decades, based on knowledge of the two toxins as well as the adhesins involved in binding the enteropathogens to host cells^{20,21}. Both LT and CT are very immunogenic and their chemistry and genetics have been studied with the aim of developing (oral) vaccines against a variety of diseases (see for example refs 21–25). Here we report the three-dimensional structure of the porcine variant of heat-labile enterotoxin and discuss its conformation with regards to properties of LT and CT.

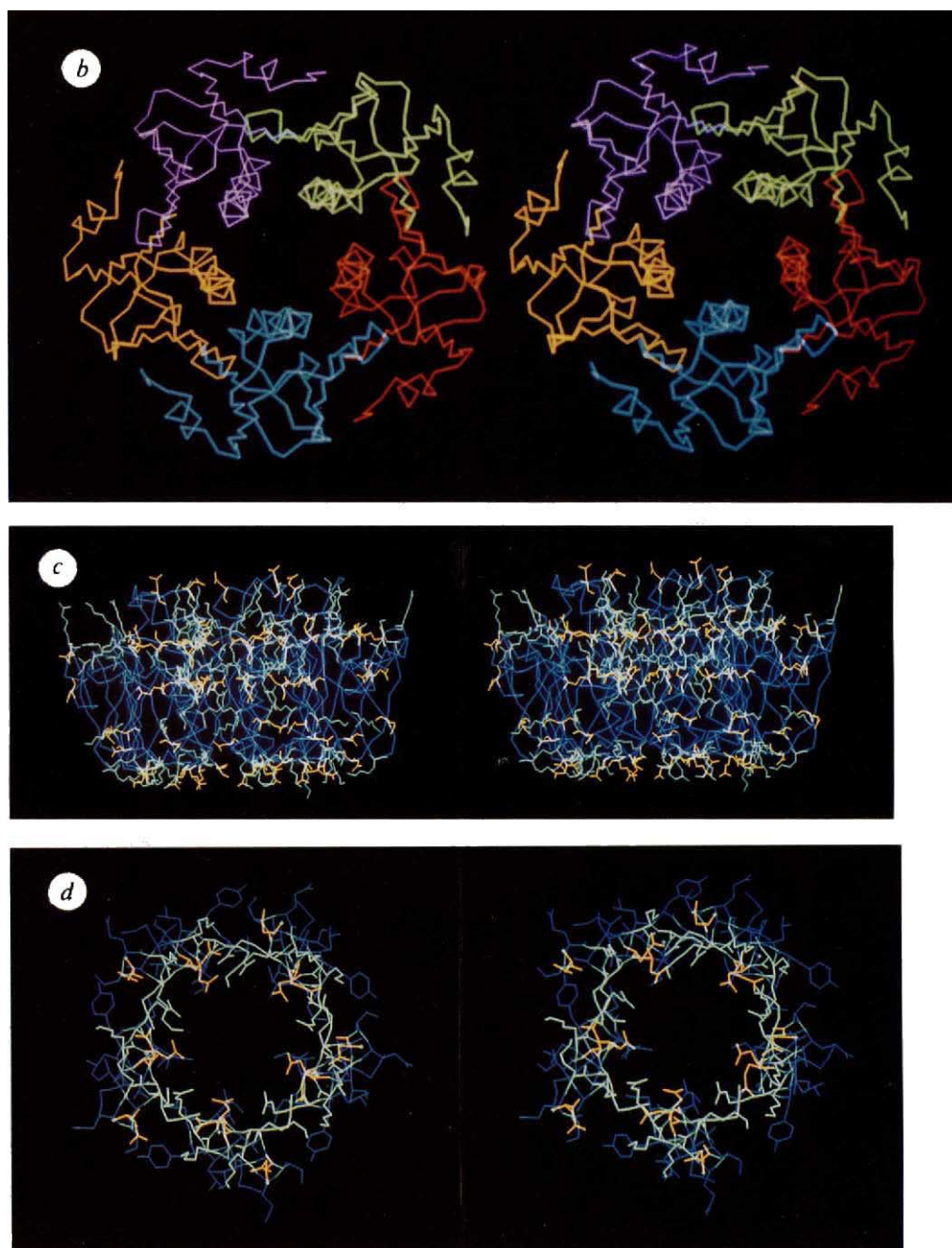
Structure determination

The structure determination of LT has been seriously hampered by nonisomorphism even between native crystals. To overcome this problem we used the same crystal twice, both for native and derivative data collection, which gave the first interpretable

Patterson map. But in crystals from a fresh batch of protein the nonisomorphism was much less and useful derivative and native data were obtained from separate crystals. Three rather poor heavy atom derivatives gave a set of starting phases, but density modification, including fivefold density averaging of the B subunits and solvent flattening, was essential for tracing the B chain. This was followed by several cycles of model building and density modification using combined multiple isomorphous replacement (MIR) and model phases, during which parts of the main chain for the A subunit could be traced. The partial AB₅ model was then refined with GROMOS molecular dynamics refinement²⁶. Finally the complete sequence of the A subunit could be determined from an electron density map with combined phase information from MIR, density modification and the refined partial model by the SIGMAA (ref. 27) procedure. Alternating cycles of model building and refinement have resulted in a model with a crystallographic *R* factor of 17.8% to a resolution of 2.3 Å and good geometry (Table 1).

The current model contains all 515 residues of the B pentamer,

FIG. 1 *a*, Schematic secondary structure diagram (projection) of a single B subunit (light grey) and the B pentamer. The secondary structure elements are α_1 : 5–9, β_1 : 15–22, β_2 : 26–30, β_3 : 37–41, β_4 : 47–50, α_2 : 59–78, β_5 : 82–88, β_6 : 94–102, deduced by the program DSSP (ref. 31). As the five B subunits have been independently refined, minor variations by one or two residues of these assignments occur in the individual B subunits. The DSSP program reports the β sheet for the pentamer as one single (circular) β sheet containing 30 'ladders' due to two hydrogen bonds between the parallel strands β_4 and β_6 within the monomer. Numbering of B subunits is anti-clockwise when viewed towards the A subunit. *b*, C α drawing of the LT B pentamer with the monomers in different colours viewed along the fivefold axis looking from the A subunit. Note the tight packing of the subunits around the central pore, formed by the long α_2 helices. *c*, The LT pentamer viewed perpendicular to the fivefold axis: a C α tracing (blue) with the negatively and positively charged residues in red and white, respectively. Note the remarkable flat 'lower' (or 'A') surface, with its high concentration of charged residues: four positively and three negatively charged residues per B monomer, 35 charges in total. At the top the exposed immunogenic region (B: 50–64) can be seen²⁹. *d*, The pentagonal helix barrel forming the central pore of the B pentamer viewed towards the A subunit. Charged residues are coloured as in *c*. Almost all 45 charged side chains have their charged functional group pointing in the same 'upwards' direction, away from the A subunit.



230 residues of the A subunit and 163 solvent molecules. The electron density is continuous for all subunits except for three regions which are not well defined in the map: (1) the N-terminal three residues of A1; (2) the C-terminal eight residues of A2, even after repeated modelling and 'omit refinement'; (3) residues A: 186–196 at the junction of A1 and A2. Electrophoresis on SDS-polyacrylamide gels (not shown) of redissolved crystals indicate that a small fraction of the protein may be nicked in this region. In addition, these residues may be inherently flexible in the intact A chain.

B subunit and pentamer

The B subunits seem to have a novel topology (Fig. 1a). The main features in the B monomer are two three-stranded anti-parallel β sheets (called I ($\beta_2, \beta_3, \beta_4$) and II ($\beta_1, \beta_5, \beta_6$), respectively), a small N-terminal helix and a large 'central' helix. The disulphide bridge (9–86) connects the N-terminal helix and strand β_5 .

The arrangement of monomers in the pentamer is interesting (Fig. 1b). Sheet I of each subunit forms a six-stranded anti-parallel β sheet with sheet II of the next subunit. The long α helices form a 'barrel' in the centre of the pentamer, creating a pore 30 Å long with a diameter ranging from ~11 Å near the A surface to ~15 Å near the 'top' of B₅. The helical barrel has a left-handed twist and is similar to that found in heavy riboflavin synthase²⁸, although in LT the cross-section of the pore is about 3 Å larger. The pentamer has a diameter of ~64 Å and a height of ~40 Å. In the 'standard view' (Fig. 3a) the 'lower' surface of the pentamer is remarkably flat. On the 'upper' side the long central helices extend above the β sheets. The exposed loop

leading into this helix comprises residues B: 50–64. As a synthetic peptide this region elicits protective antibodies against both LT and CT (ref. 29).

Interactions between B subunits occur only between nearest neighbours. These involve main-chain interactions between strands β_2 and β_6 at the subunit interface, and many contacts between side chains, like for example, the central helices that make angles of 18° with each other, ideal for side-chain interdigitation. Residue B: Ala 64 is in contact with the completely buried B: Met 31 of subunit 'N+1' (numbering as in Fig. 1), which explains why mutation of B: Ala 64 into Val interferes with pentamer formation³⁰.

The surface area of the monomer buried when the pentamer is formed is 2,360 Å², comprising 39% of the total accessible surface of the monomer³¹. This is one of the largest percentages of buried surface found in protein multimers³² and may account for the considerable stability of the pentamer (see for example ref. 33).

A subunit

The A1 fragment of LT is distinctly nonspherical with a characteristic shape: a triangle when viewed as in Figs 2b and 3a, and a V-shape when rotated by 90° (Fig. 3b). The triangle has a base of 57 Å with sides of ~35 Å. The largest dimension perpendicular to the base is 35 Å. This single domain subunit has a complicated topology with many short stretches of secondary structure, containing in total nine helices and nine β strands (Fig. 2a).

In LT, subunits A2 and A1 are linked covalently in the main chain as well as by a disulphide bridge (A: 187–199), but residues

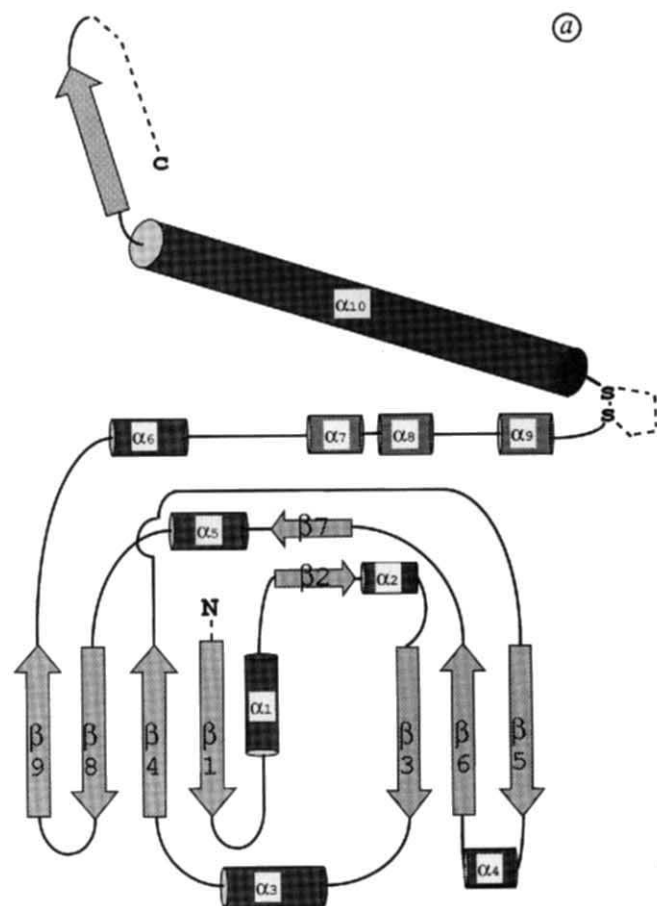


FIG. 2 The A subunit of LT. *a*, Schematic secondary structure diagram. The following secondary structure elements were assigned by the program DSSP (ref. 31): β_1 : 5–9; α_1 : 13–19; β_2 : 21–22; α_2 : 41–45; β_3 : 59–62; α_3 : 66–76; β_4 : 82–88; β_5 : 94–96; α_4 : 97–104; β_6 : 113–116; β_7 : 119–120; $\alpha_5(3_{10})$: 121–123; β_8 : 124–131; β_9 : 134–141; α_6 : 147–150; α_7 : 158–164; $\alpha_8(3_{10})$:

172–175; $\alpha_9(3_{10})$: 179–182 of A1, and, α_{10} : 200–222 of A2. *b*, Ribbon diagram of the A subunit in 'standard view', with β strands longer than four and α helices longer than five residues represented by arrows and cylindrical coils, respectively. Ribbon program by J. Priestle.

A: 186–196 could not be traced (Fig. 2a). At residue A: 200, a long 23-residue α helix starts, which continues as an extended chain. This A2 α helix is located on one of the triangular sides of A1 and it is flanked by A1 loops on both sides. There is extensive interaction with the A1 subunit, mediated by van der Waals' interactions, hydrogen bonds and one salt-bridge, between A2: Asp 225 and A1: Arg 146. The surface area buried by the A1–A2 interaction is $2 \times 1,300 \text{ \AA}^2$, 13% of the total exposed surface of A1 and $\sim 26\%$ of that of A2.

AB₅ complex

The interaction between the A subunit and the B₅ pentamer is almost entirely mediated by A2 (Fig. 3), which agrees well with the results of previous studies¹³. Remarkably, the extended A2 chain 222–228 continues all the way through the central B₅ pore. At the top of the pentamer the chain turns back into the pore although the last eight residues are not well defined. The A2 subunit has different interactions with each of the five B subunits. The surface buried by the interaction of A2 with B₅ is $\sim 2 \times 1,700 \text{ \AA}^2$. This implies that over 50% of the A2 surface is buried in the AB₅ complex. The central B helices line the central wall with charged residues: Asp 59, Lys 62, Lys 63, Glu 66, Arg 67, Lys 69, Asp 70 and Arg 73; a total of not less than 25 positive and 15 negative charges (Fig. 1d). The A2 hairpin buried in this pore contains five negative and two positive charges, not entirely compensating for the overall charge of the pore.

In contrast to A2, there are very few direct interactions between the enzymatic A1 fragment and the B pentamer. Only two loops of A1 (Fig. 3b), one on each side of the A2 helix, interact with the second, third and fourth B subunits (B(2), B(3) and B(4), respectively). The A1–B(2) contact is just the salt-bridge A1: Arg 33–B(2): Glu 79. At the other side of the A2 helix in the very hydrophilic loop A1: 147–151 (Asp–Arg–Tyr–Tyr–Arg) every residue except A1: Tyr 150 has interactions with B(3) or B(4).

Ganglioside binding and translocation

Tryptophan B: 88 of LT and CT has been implicated as part of

the binding site of the G_{M1} oligosaccharide moiety^{12,34,35}. In the structure B: Trp 88 is found at the bottom of a small cavity (Fig. 4), surrounded by several loops, including one containing B: Gly 33 from subunit 'N + 1'. Mutation of B: Gly 33 into Asn causes loss of G_{M1} binding activity³⁶. B: Gly 33 takes part in a β turn and any mutation at that position is likely to alter the local backbone conformation besides adding a side chain. The proximity between atoms of B: Trp 88 (C ϵ) and B: Gly 33 (N) (4.8 $\text{\AA}$), indicates that this cavity is indeed part of the ganglioside binding site. It implies also that G_{M1} binding requires two B subunits, a proposal made by several authors^{30,34}. It is possible that the terminal galactose of G_{M1} interacts with B: Trp 88, analogous to some other sugar-binding proteins³⁷.

Molecular events occurring on G_{M1} binding are essential for toxin action. The B pentamer neither enters the bilayer^{38,39} nor undergoes major conformational changes^{40–42}, and binds with its plane parallel to the membrane⁴². A crucial question is whether binding of LT and CT occurs with the A subunit directed away from the membrane or towards it.

There is evidence for binding with A towards the membrane^{39,41,42}, and some for the alternative position^{13,38}. In the structure of LT presented here, the surface of the B pentamer on the side where the A subunit is situated is not only very flat, but also highly charged (Fig. 1c) and this surface could potentially interact with the charged groups on the membrane. On the A subunit no such highly charged areas were found, nor were large patches of exposed hydrophobic residues. In the B subunits the putative binding site for G_{M1} near Trp 88 is about 23 $\text{\AA}$ from this flat charged surface. Because G_{M1} extends $\sim 25 \text{ \AA}$ from the membrane surface⁴³, it is possible that LT/CT may bind with the A subunit inserted into the membrane, where the oligosaccharide chains of G_{M1} extend like cables from the membrane surface all the way to the top of the B subunits.

NAD binding and relation to exotoxin A

The A subunit of LT does not contain the classical NAD-binding $\beta\alpha\beta$ -fold. The structure of one other ADP-ribosylating toxin is known: exotoxin A (ETA) from *Pseudomonas aeruginosa*

FIG. 3 Three-dimensional structure of heat-labile enterotoxin. Stereo C α picture of the LT AB₅ complex. The A2 subunit is shown in thick lines. CA1, last residue of the A1 chain, which is well defined in the electron density map (A1: 187); NA2, first well-defined residue of A2 (197). The structure of LT has been determined in an 'inactive' form, which requires further nicking and disulphide reduction for cellular activity^{11–13}. As the 'nicking' between residues A1:192 and A2:195 appears to take place at least 25 $\text{\AA}$ from the essential active-site residues, it is possible that this modification is required not for activating an 'inactive' active site, but for correct interaction with one or more G proteins. a, 'Standard view' of LT; orientation with 'triangular view' of A1. All relative orientations refer to this standard setting. Positions of A1: Glu 112 and B: Trp 88 are indicated. From the positions of the C termini of the B subunits it is clear why B subunits elongated at the C terminus do not form AB₅ complexes⁶². b, Orientation with 'cone' view of A1, rotated around the fivefold axis by 90° with respect to the triangular view in a. Note that the contacts between A1 and B₅ are very limited indeed.

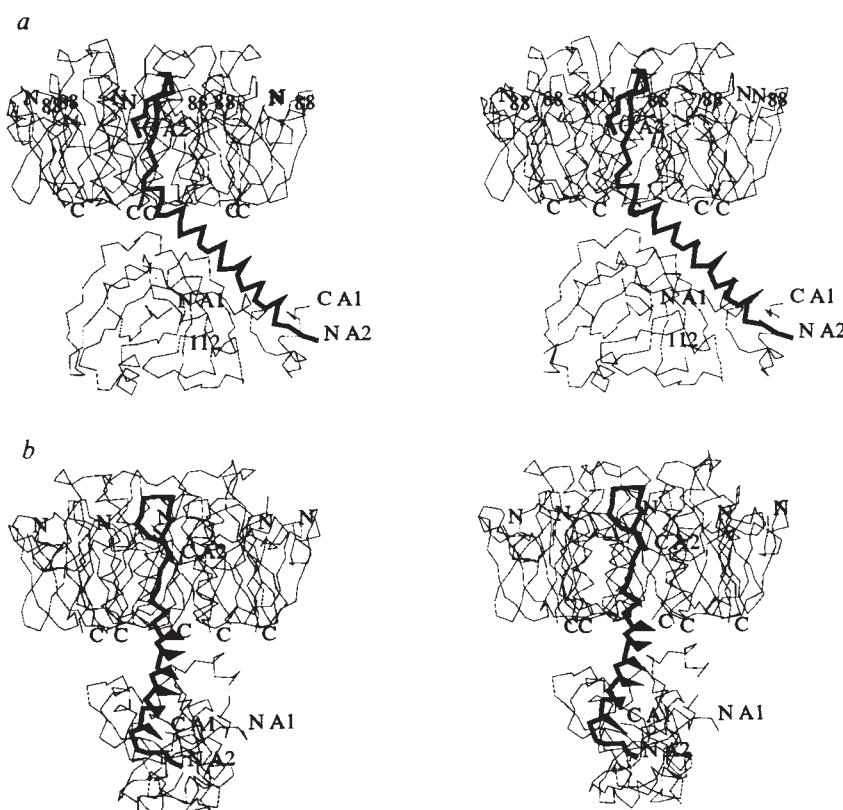
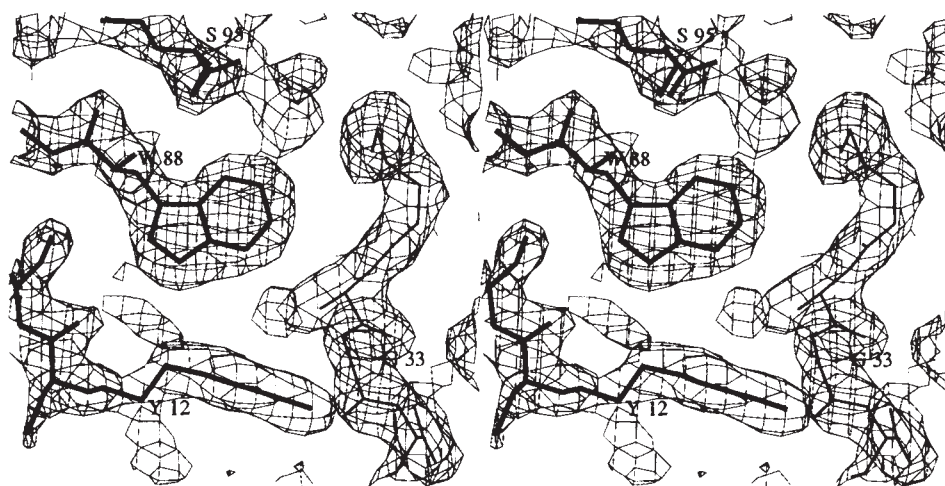


FIG. 4 Electron density map showing the surroundings of B: Trp 88 forming part of the putative G_{M1} binding site viewed from the top of the B subunit towards the A subunit (see Fig. 3a). Residues of subunit B(1) are shown in thick lines. In thinner lines part of the β hairpin containing B: Gly 33 of subunit B(2) is shown. All residues shown are identical between CT and LT, except B: Ser 95, which is an Ala in CT. The density shown is a $(2m|F_o| - D|F_c|)$ SIGMA-weighted²⁷ map of the refined model at a resolution of 2.3 Å contoured at one times the standard deviation.



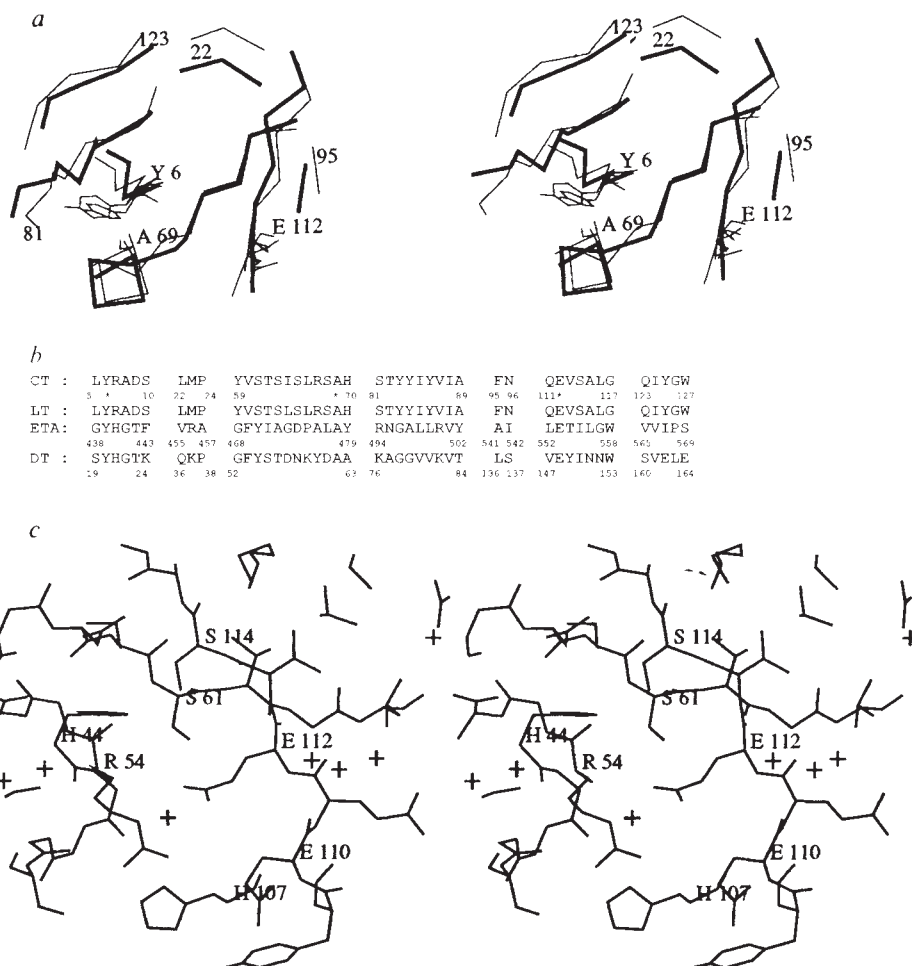
(ref. 44). This single-chain protein has three distinct domains, with very different membrane translocation properties from LT/CT. ETA ADP-ribosylates a diptamide residue, which is a modified histidine, in elongation factor-2 (EF-2), as does diphtheria toxin, with which it shares some sequence homology in the enzymatic domain^{45,46}.

The structural superposition of the enzymatic domain of ETA on the LT A subunit shows significant similarity (Fig. 5a): 44 residues can be superimposed with an r.m.s. difference of 1.5 Å on C_α coordinates. The equivalent residues roughly coincide with helix α_3 and strands β_1 to β_6 , which form the central part of both enzymes. There is virtually no sequence homology; only

three of the 44 residues are identical (Fig. 5b): A1: Tyr 6, Ala 69 and Glu 112 in LT. The latter amino acid corresponds to Glu 553 in ETA, which is an active-site residue because it can be covalently linked to NAD and loses activity after mutation to Asp (refs 46, 47). Residue A1: Glu 112 in LT has been shown to be important for activity by mutagenesis studies⁴⁸.

Residue A1: Glu 112 is hydrogen-bonded to the hydroxyl group of A1: Ser 61 (Fig. 5c), which was also shown to be essential for activity by mutagenesis studies⁴⁹. The structure reveals several other groups in the immediate neighbourhood that may be important for catalysis or recognition: His 44 and 107, Ser 114 and Arg 54, the latter forming a salt-bridge to

FIG. 5 The LT active site. *a*, Superposition of LT A on the enzymatic domain of exotoxin A from *P. Aeruginosa* (coordinates provided by D. McKay) in the orientation of Fig. 3a. Equivalent C_α residues are shown in thick lines for LT (with residue numbers) and thin lines for ETA. Identical residues in LT and ETA (A: Tyr 6, Ala 69, Glu 112) are shown completely. Note that the LT A chain was traced without reference to the structure of exotoxin A. *b*, Sequence alignment of 44 structurally equivalent residues in LT⁷ and ETA⁶³ as well as the sequentially similar diphtheria toxin (DT)^{64,65} and CT²⁰. Alignment of ETA and DT sequences slightly extrapolated from those in refs 46 and 47. Asterisks indicate identical residues between ETA and LT. *c*, Stereo diagram of the active site in the A subunit of LT. The orientation is as in Fig. 3b. The labelled residues are discussed in the text.



A1: Glu 112. These residues are in an elongated crevice 'under' strand β_3 , as in the standard view of Fig. 3a.

Another interesting residue is A1: Arg 146, the site of auto-ADP-ribosylation in LT (ref. 50), although mutation to a Gly did not cause a change in activity⁵¹. The guanidinium group of A1: Arg 146 is 30 Å from A1: Glu 112, suggesting that LT probably carries out a 'cross' instead of a 'self' auto-ADP-ribosylation.

Residue A1: Ser 61 superimposes on Tyr 470 in ETA. In ETA this residue can be mutated to Phe without change in activity⁵², whereas there are no structurally equivalent residues for the other residues surrounding A1: Glu 112 mentioned above, except A1: Ser 114, which is an Ile in ETA. The active sites of the two proteins are very different, which agrees with the difference between the two affinity constants for NAD (ref. 53). The structural similarity is interesting, possibly suggesting a

distant evolutionary relationship. It is possible that LT/CT and ETA share a new NAD-binding fold that is conserved widely among ADP-ribosylating proteins, such as diphtheria toxin, pertussis toxin and heat-labile toxin II (ref. 54), and perhaps also in the intracellular ADP-ribosylating proteins^{1,2}.

Conclusion

The structure of LT exhibits an unusual AB₅ architecture and allows localization of part of the receptor binding site in the B subunit and of the active site in the A subunit. Detailed characterization of the G_{M1}-oligosaccharide, NAD and arginine binding sites should enable compounds to be designed based on protein structure which block toxin action and may be used as drugs. The structure will also be of considerable value for engineering potential vaccines, as candidate surface loops for epitope insertion can now be identified easily. □

Received 28 February; accepted 22 April 1991.

- Foster, J. W. & Kinney, D. M. *CRC crit. Rev. Microbiol.* **11**, 273-298 (1985).
- Moss, J. & Vaughan, M. *Adv. Enzym.* **61**, 303-379 (1988).
- van Heyningen, S. in *Current Topics in Membranes and Transport* Vol. 18, 445-471 (Academic, New York, 1983).
- Finkelstein, R. A. in *Immunochemical and Molecular Genetic Analysis of Bacterial Pathogens* (eds Owen, P. & Foster, T. J.) 85-102 (Elsevier New York, 1988).
- Geary, S. J., Marchlewicz, B. A. & Finkelstein, R. A. *Infect. Immunity* **36**, 215-220 (1982).
- Dallas, W. S. & Falkow, S. *Nature* **288**, 499-501 (1980).
- Yamamoto, T., Gojobori, T. & Yokota, T. *J. Bact.* **169**, 1352-1357 (1987).
- Hirst, T. R., Sanchez, J., Kaper, J. B., Hardy, S. J. S. & Holmgren, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7752-7756 (1984).
- Hofstra, H. & Witholt, B. *J. biol. Chem.* **259**, 15182-15187 (1984).
- Spicer, E. K. & Noble, J. A. *J. biol. Chem.* **257**, 5716-5721 (1982).
- Mekalanos, J. J., Collier, R. J. & Romig, W. R. *J. biol. Chem.* **254**, 5855-5861 (1979).
- Moss, J., Osborne, J. C. Jr, Fishman, P. H., Nakaya, S. & Robertson, D. C. *J. biol. Chem.* **256**, 12861-12865 (1981).
- Gill, D. M. *Biochemistry* **15**, 1242-1248 (1976).
- Gill, D. M. & Richardson, S. H. *J. infect. Dis.* **141**, 64-70 (1980).
- Cassel, D. & Pfeuffer, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2669-2673 (1978).
- Moss, J. & Vaughan, M. *J. biol. Chem.* **252**, 2455-2457 (1977).
- Kahn, R. A. & Gilman, A. G. *J. biol. Chem.* **261**, 7906-7911 (1986).
- Gill, D. M. & Coburn, J. *Biochemistry* **26**, 6364-6371 (1987).
- Peterson, J. W. & Ochoa, L. G. *Science* **245**, 857-859 (1989).
- Mekalanos, J. J. et al. *Nature* **306**, 551-557 (1983).
- Clemens, J. D. et al. *Lancet* **II**, 124-127 (1986).
- Schödel, F. & Will, H. *Infect. Immunity* **57**, 1347-1350 (1989).
- Sanchez, J., Svennerholm, A.-M. & Holmgren, J. *FEBS Lett.* **241**, 110-114 (1988).
- Dertzbaugh, M. T., Peterson, D. L. & Macrina, F. L. *Infect. Immunity* **58**, 70-79 (1990).
- Czerkinsky, C., Russell, M. W., Lycke, N., Lindblad, M. & Holmgren, J. *Infect. Immunity* **57**, 1072-1077 (1989).
- Fujinaga, M., Gros, P. & van Gunsteren, W. F. *J. appl. Crystallogr.* **22**, 1-8 (1989).
- Read, R. J. *Acta crystallogr.* **A42**, 140-149 (1986).
- Ladenstein, R. et al. *J. molec. Biol.* **203**, 1045-1070 (1988).
- Jacob, C. O., Leitner, M., Zamir, A., Salomon, D. & Arnon, R. *EMBO J.* **4**, 3339-3343 (1985).
- Iida, T. et al. *J. biol. Chem.* **264**, 14065-14070 (1989).
- Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577-2637 (1983).
- Janin, J., Miller, S. & Chothia, C. *J. molec. Biol.* **204**, 155-164 (1988).
- Goins, B. & Freire, E. *Biochemistry* **27**, 2046-2052 (1988).
- De Wolf, M. J. S., Frickin, M. & Kohn, L. D. *J. biol. Chem.* **256**, 5489-5496 (1981).
- Ludwig, D. S., Holmes, R. K. & Schoolnik, G. K. *J. biol. Chem.* **260**, 12528-12534 (1985).
- Tsuji, T., Honda, T., Miwatani, T., Wakabayashi, S. & Matsubara, H. *J. biol. Chem.* **260**, 8552-8558 (1985).
- Montfort, W. et al. *J. biol. Chem.* **262**, 5398-5403 (1987).
- Tomasi, M. & Montecucco, C. *J. biol. Chem.* **256**, 11177-11181 (1981).
- Wisniewski, B. J. & Bramhall, J. S. *Nature* **289**, 319-321 (1981).
- Surewicz, W. K., Leddy, J. J. & Mantsch, H. H. *Biochemistry* **29**, 8106-8111 (1990).
- Dwyer, J. D. & Bloomfield, V. A. *Biochemistry* **21**, 3227-3231 (1982).
- Ribi, H. O., Ludwig, D. S., Mercer, K. L., Schoolnik, G. K. & Kornberg, R. D. *Science* **239**, 1272-1276 (1988).
- McDaniel, R. V. & McIntosh, T. J. *Biophys. J.* **49**, 94-96 (1986).
- Allured, V. S., Collier, R. J., Carroll, S. F. & McKay, D. B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1320-1324 (1986).
- Brandhuber, B. J., Allured, V. S., Falbel, T. G. & McKay, D. B. *Proteins* **3**, 146-154 (1988).
- Carroll, S. F. & Collier, R. J. *J. biol. Chem.* **262**, 8707-8711 (1987).
- Douglas, C. M. & Collier, R. J. *J. Bact.* **169**, 4967-4971 (1987).
- Tsuji, T. et al. *J. biol. Chem.* **265**, 22520-22525 (1990).
- Harford, S., Dykes, C. W., Hobden, A. N., Read, M. J. & Halliday, I. J. *Eur. J. Biochem.* **183**, 311-316 (1989).
- Lai, C.-Y., Xia, Q.-C. & Salotra, P. T. *Biochem. biophys. Res. Commun.* **116**, 341-348 (1983).
- Okamoto, K. et al. *J. Bact.* **170**, 2208-2211 (1988).
- Lukac, M. & Collier, R. J. *Biochemistry* **27**, 7629-7632 (1988).
- Galloway, T. S. & van Heyningen, S. *Biochem. J.* **244**, 225-230 (1987).
- Pickett, C. L., Twiddy, E. M., Coker, C. & Holmes, P. K. *J. Bact.* **171**, 4945-4952 (1989).
- Pronk, S. E. et al. *J. biol. Chem.* **260**, 13580-13584 (1985).
- Messerschmidt, A. & Pflugrath, J. W. *J. appl. Crystallogr.* **20**, 306-315 (1987).
- Kabsch, W. *J. appl. Crystallogr.* **21**, 916-924 (1988).
- Leslie, A. G. W. *Acta crystallogr.* **A43**, 134-136 (1987).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
- Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819-822 (1986).
- Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489-501 (1987).
- Sandkvist, M., Hirst, T. R. & Bagdasarian, M. *J. Bact.* **169**, 4570-4576 (1987).
- Gray, G. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2645-2649 (1984).
- Greenfield, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6853-6857 (1983).
- Ratti, G., Rappouli, R. & Giannini, G. *Nucleic Acids Res.* **11**, 6589-6595 (1983).

ACKNOWLEDGEMENTS. We thank all members of the Groningen Protein Crystallography Group for help and support during this project, Z. Kangjing and H. Groendijk for assistance in heavy atom derivative search, J. Kingma for protein purification and D. McKay for the coordinates of exotoxin A. This research was supported by the Netherlands Foundation for Chemical Research (SON) and the Netherlands Organisation for Scientific Research (NWO).

Time-resolved holography

A. Rebane* & Jack Feinberg

Department of Physics, University of Southern California, Los Angeles, California 90089-0484, USA

IN a conventional hologram, a photographic film records the interference pattern of monochromatic light, scattered from the object to be imaged, with a reference beam of unscattered light. Illumination of the developed film with a replica of the reference beam then creates a virtual image of the original object. Here we show how a molecular resonance can be used to record an interference pattern between light signals that arrive at different times, and with this technique create a hologram with time resolution. Using a timed reference pulse as a 'light shutter', we can record holographic images selectively, according to the time taken by light travelling from the object to the hologram. We use this method to image an object behind a semi-opaque screen, and indicate how a similar method could be used to inspect objects embedded in a dense scattering medium. Ultimately, this technique might be applied to the medical imaging of tumours.

Consider two optical pulses incident on a holographic recording medium as shown in the inset of Fig. 1. Let pulse F arrive T_0 seconds before pulse G . If we use photographic film as the recording medium, then an interference pattern will be recorded only if the pulses overlap with each other at least partially in time. But suppose we replace the photographic film by a bank of resonators such as atoms, with each atom tuned to a slightly different optical frequency ω_j . In this case, the interference pattern of the two optical pulses F and G can be recorded even if the pulses are never present in the material at the same time¹.

If the optical pulse F is sufficiently brief, then its frequency spread will be wide enough to excite all of the atoms, much as a brief kick to a piano will excite all of the piano's strings. Each atom will continue to ring at its own natural frequency ω_j . After a delay time T_0 , the second light pulse G arrives, and it will transfer energy either into or out of the j th atom depending on the relative phase between the optical electric field of G and the phase $\phi_j = \omega_j T_0$ of the still-ringing atom. Because each atom's phase depends on its particular resonant frequency ω_j , a given time delay T_0 between the two incident light pulses will produce a unique pattern of excited and unexcited atoms in frequency space.

After interacting with both of the light pulses F and G , let the absorption of the j th atom be permanently altered by an amount proportional to the atom's final energy (by a mechanism described below). If the now altered atoms are illuminated by a replica of pulse F , they will absorb and re-radiate light, again at the frequencies ω_j appropriate to each atom. At first, the phases of the re-radiated light will be incoherent, but after a time of exactly T_0 , the phases of the different frequencies come together to reproduce coherently a duplicate of the pulse G (ref. 1). It can be shown², however, that application of a replica of pulse G to the altered atoms causes a re-radiation of light that loses rather than gains phase coherence with the passage of time. Unlike conventional off-axis holography, where either light beam can be used to reconstruct the other, the bank of atoms here records the direction of time's arrow.

In fact, as we describe below, it is possible to reproduce an image of the scattered light that strikes the hologram before the reference pulse arrives. To produce such an image, the hologram is read out by a replica of pulse F applied in the opposite direction to the original pulse: reversing the spatial orientation of the readout pulse reverses the relative phase relation of the re-radiated light, and thereby leads to the recreation of the 'before' rather than 'after' scattered light.

Now consider illuminating an entire scene with a pulsed laser. Record the reflected light G using a bank of tuned resonators, and let a reference pulse F also illuminate the resonators, as described above. If the reference pulse arrives at the resonators before any of the reflected light, then the entire scene can be recalled by simply reading with another reference pulse. But if the initial reference pulse is delayed so that it arrives after some of the reflected light, then reading (with another reference pulse) will only recreate the parts of the image that arrived after that first reference pulse, which are the parts of the scene that were located farthest from the resonators. For example, if the scene consisted of the street view of a bookshop, then, with a suitably timed reference pulse, the reconstructed image would show the books on display deep inside the store, and would not show the reflections off the shop's front window.

We constructed a bank of narrow-band resonators by doping a solid block of polystyrene with the organic dye protoporphyrin²⁻⁶. Each dye molecule acts as a lightly damped resonator, but because of inhomogeneities in the plastic matrix, each dye molecule has a slightly different resonant frequency. When illuminated by narrow-band light of frequency ω , the molecules that happen to be in resonance with the light become excited. A fraction of these excited molecules subsequently relaxes into a metastable state (thought to be a tautomerization of the original molecule). Once in this transformed state the molecule's absorption is shifted to a completely different spectral region, so a narrow 'hole' is burned into the sample's absorption spectrum at the frequency ω . This spectral hole remains as long as the sample is kept cold. If cooled to a temperature of 2 K, the phase-relaxation time, T_2 , of the molecule's upper level becomes quite long ($T_2 \approx 1$ ns), so that the absorption hole has a narrow homogeneous spectral width of $\Delta\omega = 1/(\pi T_2) \approx 0.01$ cm⁻¹. The polystyrene matrix causes the net absorption spectrum of all of the molecules to form an inhomogeneously broadened band extending over a range of 200 cm⁻¹. Consequently, this material resembles a bank of 200/0.01 = 20,000 narrow-bands of resonators and we use this 'spectral-hole-burning' material to record and store the spectral (and, as we show below, spatial) contents of an incident light beam.

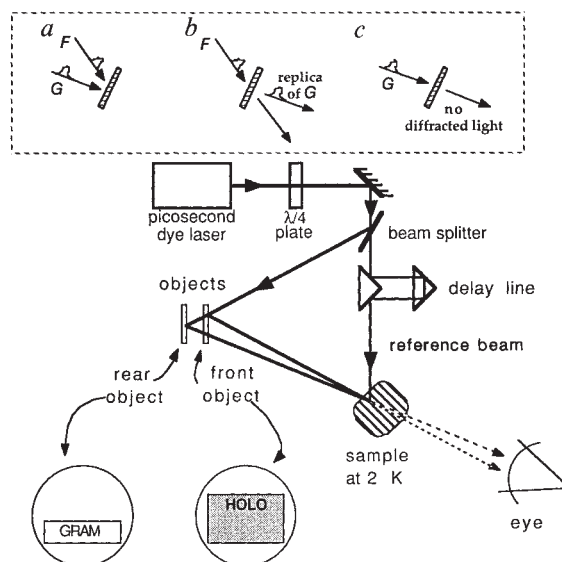


FIG. 1 Main figure: experimental set up. The front object is a frosted slide with the letters 'HOLO' pasted on the front. The rear object is the letters 'GRAM' pasted on the back of the same slide. Inset: a, Writing the hologram with two light pulses separated in time. b, Reading with the earlier pulse F recreates the later pulse G . c, Reading with the later pulse G does not produce any diffracted pulse.

* Permanent address: Institute of Physics, Estonia Academy of Science, 142 Riia Street, Tartu, Estonia, USSR

In our experiments, we used a 3-mm-thick block of polystyrene doped with protoporphyrin at a concentration of $10^{-3} \text{ mol l}^{-1}$. The useful sample area was 4 cm^2 . The sample was immersed in liquid helium, and the helium vapours were pumped to reduce the temperature to 2 K. The peak of the broad absorption feature was at $\lambda = 621 \text{ nm}$, at which the optical density was 1.6. The light source was a continuous-wave modelocked Nd:YAG laser (Coherent Antares 76s) which synchronously pumped a tunable dye laser (Coherent 701) to produce pulses having an intensity width of 8 ps full width at half maximum (FWHM). (These pulses are not transform-limited; they have a coherence width of only 0.5 ps.) The repetition rate of the laser pulses was 76 MHz. A beamsplitter divided the beam from the picosecond dye laser into a reference beam and a separate beam to illuminate the various objects in the scene, as shown in Fig. 1. The reference beam was expanded by a telescope to illuminate the entire polystyrene block. Scattered light from the illuminated objects simply propagated to the polystyrene block with no intervening lens. The angle between the reference beam and the image-bearing beam was $\sim 14^\circ$.

The recorded scene consisted of two objects. The nearby object was a 1.0-mm-thick glass slide with the letters 'HOLO' attached to its front surface. The distant object was a white paper screen carrying the letters 'GRAM', which was pressed against the back of the transparent slide, so the separation between the two objects was $\sim 1 \text{ mm}$. To increase the amount of light scattered by the slide, its front surface was coated with a frosting aerosol spray (New York Bronze Powder Co.) The slide was illuminated from the front. Of the scattered light from the slide reaching the polystyrene block, $\sim 80\%$ came from the front sprayed surface of the slide, and only $\sim 20\%$ came from the rear surface. Viewing the laser-illuminated slide by eye from

the position of the polystyrene block, one could clearly read the words HOLO, but the intense glare from the front surface almost completely obscured the letters GRAM located near the back of the slide.

Light from the laser reached the front of the slide 5 ps before it reached the back of the slide. Consequently, light from the front of the slide reached the storage medium 10 ps (twice the glass travel-through time) before light from the back of the slide. In the first experiment a hologram was recorded with the reference beam timed to arrive before both of these object waves. In the second experiment a hologram was recorded with the reference beam carefully timed to arrive within the 10-ps interval between the two object waves.

The reference beam and the object beam each had an average intensity of 0.2 mW cm^{-2} at the location of the storage medium. An exposure time of 2–3 minutes was needed to record a hologram with a fluence of $70\text{--}100 \text{ mJ cm}^{-2}$, corresponding to 10^{10} identical pairs of laser pulses. To write the first hologram we tuned the laser wavelength to the absorption maximum of the medium at 621 nm ($19,950 \text{ cm}^{-1}$). Because the absorption spectrum is 200 cm^{-1} wide and the dye-laser pulses have a spectral bandwidth of only $\sim 30 \text{ cm}^{-1}$, we could change the wavelength of the dye laser by $\sim 30 \text{ cm}^{-1}$ and then record a new hologram in a 'fresh' spectral region of the storage medium without affecting any previously stored holograms. The different holograms written at different centre wavelengths were later selectively read out by simply adjusting the centre wavelength of the reading dye-laser pulse. In the absence of all illumination these holograms lasted for as long as the sample was kept cold.

We read the holograms by blocking the light coming from the objects and illuminating the hologram with only the reference beam. Figure 2a shows the reconstructed image when the

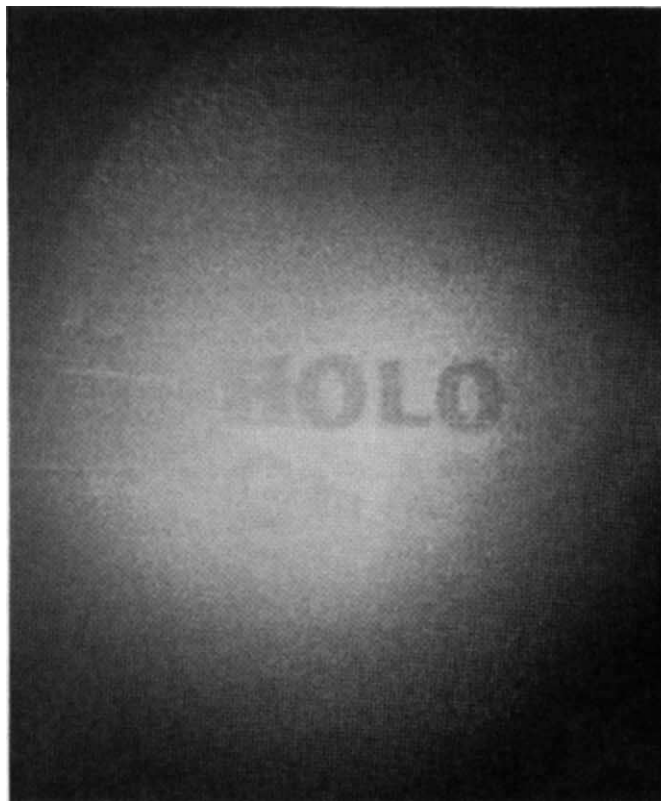


FIG. 2 *a*, Holographic reconstruction of both objects. The reference pulse was set to arrive before light from either of the objects. The glare from the frosted glass in front with the word HOLO obscures the object GRAM behind it. *b*, In this hologram the reference pulse was delayed to arrive after

light from the front object (frosted glass) but before light from the rear object. The front object is no longer reconstructed, whereas the rear object (GRAM) is now plainly visible.

hologram was recorded using a reference beam that arrived a few picoseconds before any of the light from the glass slide. The letters HOLO on the front of the slide are plainly visible, but the glare from the front of the slide almost obscures the letters GRAM near the back of the slide. Figure 2b shows the reconstructed image when the reference beam pulse was carefully set to arrive after light from the distant object but before light from the nearby object. Now only the distant object was reconstructed, and the nearby object was eliminated. We emphasize that the light from these objects need not arrive at the polystyrene block at the same time as light in the reference pulse for the hologram to be recorded. In fact, because the coherence time of the light pulses was only 0.5 ps, and because we set the reference beam to arrive at the storage medium ~ 5 ps before the light from the back of the slide, the reference and object beams could not have produced a conventional intensity interference pattern in the storage medium. We note that the maximum time delay permitted between the object and reference beams was 10^3 ps and is set by the phase-decay time T_2 of the sample.

In the experiment reported here we could selectively reconstruct those objects that sent light to the hologram after the reference pulse had arrived. For some applications, however, it is desirable to do the opposite and recreate the light that arrived before the reference pulse. For example, consider the problem of imaging an object that is embedded in a scattering medium, such as a tumour embedded in breast tissue. Illuminate the

tissue from behind with a short laser pulse. Light transmitted through the tissue without any scattering will emerge before light that has been multiply scattered by the tissue⁷. Because the eye records all of this light, and because the scattered light overwhelms the unscattered light, the tumour remains unseen. But if light that arrived earlier at the eye could be selectively enhanced, than a shadowgram of the tumour would become visible.

This selection of early light over late light can be accomplished by simply altering the direction of the readout beam used in the experiments above. Instead of using a readout beam in the same direction as the reference beam, one should use a readout beam that is directed exactly opposite to the direction of the original reference beam. This is the 'four-wave mixing' geometry of traditional phase-conjugation experiments⁸, but with a spectral-hole-burning material now only the light that arrived before the reference beam is holographically reconstructed in the final image. In this way light scattered from back-lighted tissue could be eliminated, while light travelling directly through the tissue could be preserved, and would form a shadowgram of features embedded in the tissue. We caution that this scheme requires a very large dynamic range for the hologram, because in thick tissue the scattered light can be much more intense than the unscattered light. We estimate that our present spectral-hole-burning medium has a dynamic range limited to $\sim 10^3$. Experiments to demonstrate such selective imaging are in progress. $\square$

Received 11 February; accepted 29 April 1991.

1. Longuet-Higgins, H. C. *Nature* **217**, 104 (1968).
2. Saari, P., Kaarli, R. & Rebane, A. *J. opt. Soc. Am.* **B3**, 527-533 (1986).
3. Mossberg, T. W. *Opt. Lett.* **7**, 77-79 (1982).
4. Rebane, A., Kaarli, R., Saari, P., Anijalg, A. & Timpmann, K. *Opt. Comm.* **47**, 173-176 (1983).
5. Rebane, A. *Opt. Comm.* **65**, 175-178 (1988).

6. Moerner, W. E. *Persistent Spectral Hole Burning: Science and Applications* (Springer, Berlin, Heidelberg, 1988).

7. Yoo, K. M. & Alfano, R. R. *Opt. Lett.* **15**, 320-322 (1990).

8. Hellwarth, R. W. *J. opt. Soc. Am.* **67**, 1-3 (1977).

ACKNOWLEDGEMENTS. J. F. is grateful for support from the US Air Force Office and their University Research Initiative.

Effects of pressure and stress on C₆₀ fullerite to 20 GPa

Steven J. Duclos*, Keith Brister†, R. C. Haddon*, A. R. Kortan* & F. A. Thiel*

* AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

† Cornell High Energy Synchrotron Source, Cornell University, Ithaca, New York 14853, USA

BULK C₆₀ (fullerite) is a solid composed of extremely hard pseudospherical molecules bonded by weak van der Waals interactions. We expect that pressurization will probe this diversity as the intermolecular distance is reduced. Here we present experimental measurements of the room-temperature pressure-volume equation of state of C₆₀ fullerite to pressures of 20 GPa. The face-centred cubic phase remains stable under hydrostatic compression to this pressure, with an atmospheric-pressure isothermal bulk modulus of $K_0 = 18.1 \pm 1.8$ GPa and $dK_0/dP = 5.7 \pm 0.6$, characteristic of isotropic van der Waals intermolecular bonding. X-ray diffraction intensity changes are consistent with the compressibility of the C₆₀ molecule being significantly lower than that of the bulk fullerite solid. Under non-hydrostatic compression, a transition to a crystallographic structure of lower symmetry is observed at 16 ± 1 GPa.

The recent discoveries that bulk quantities of the C₆₀ molecule can be crystallized¹ and intercalated to give metallic² and superconducting³ materials have prompted investigations into the physical properties of both the C₆₀ molecule and the solid fullerite. The molecular solid is characterized by strong sp^2 bonding within the C₆₀ molecules⁴ and weak van der Waals bonding between the molecules. Under the application of pressure,

this intermolecular bonding is expected to result in a relatively soft equation of state (EOS) at low compression. Experiments on graphite indicate a very low compressibility of the sp^2 in-plane carbon bonds of hexagonal graphite, similar to that found for the sp^3 bonds in diamond⁵. Unlike the layered graphite structure, in which the sp^2 bonds lie in two dimensions, the fullerite structure has sp^2 bonds that curve into three dimensions. This difference is expected to manifest itself in the bulk properties of the solid. As the intermolecular distances are reduced with increasing pressure, the bulk compressibility should approach that of the C₆₀ molecule and should allow studies of the stability of the molecule itself. Recent theoretical treatments have predicted fullerite bulk moduli at 20 GPa as diverse as 640 GPa (ref. 6) and 180 GPa (D. Tomanek, personal communication). We have measured the pressure-volume EOS of solid C₆₀ to 20 GPa using *in situ* X-ray diffraction from a diamond-anvil cell.

All samples used in this study were from the same preparation of powdered single crystals of C₆₀ grown by 600 °C sublimation of chromatographed graphitic soot generated by the method of Krätschmer *et al.*¹. Single-crystal X-ray diffraction at atmospheric pressure⁷ shows these samples to be face-centred cubic (f.c.c.) with a lattice constant $a_0 = 14.198$ Å. The samples were loaded into gaskets of stainless or spring steel which had been pre-indented to 75 μ m thickness. The results of three experiments are summarized in Table 1. Pressures were measured using the ruby luminescence method⁸. Table 1 shows the maximum pressure at which each experiment could be considered hydrostatic. Above this maximum pressure the R_1 and R_2 peaks of the ruby luminescence were not clearly resolved. Experiment 1 was hydrostatic throughout, whereas experiment 3 became nonhydrostatic at 4 GPa. Pressure gradients within the sample chamber were monitored in experiments 1 and 2

TABLE 1 Conditions for the three high-pressure experiments on C_{60}

Expt	Pressure medium*	Sample diameter (μm)	Maximum pressure (GPa)	Max. hydrostatic pressure (GPa)
1	4:1 ME	200	20	20
2	4:1 ME	150	27	10
3	none	200	24	4

* 4:1 ME denotes a 4:1 mixture of methanol-ethanol.

using several ruby chips distributed across the sample.

Data were collected on the B-1 station of the National High Pressure Facility at the Cornell High Energy Synchrotron Source (CHESS). The f.c.c. lattice constant a was measured using energy-dispersive X-ray diffraction (EDXD) with a high-resolution intrinsic germanium detector. A tapered pinhole $60\ \mu\text{m}$ in diameter was used as an aperture to illuminate the sample with X-rays. Diffraction angles of $2\theta = 4.8^\circ$ were used, giving diffracted energies of 18–48 keV. As the lattice constants of the sample and steel gaskets are so different, the diffraction peaks of the latter are completely separated in energy and do not interfere with the analysis. Care was taken to ensure that no Laue spots from the diamond interfered with the powder diffraction from the sample. A typical diffraction pattern is shown in Fig. 1.

Figure 2 shows the room-temperature EOS of f.c.c. C_{60} . The density of fullerite at 20 GPa is 69% that of diamond at this pressure, and is stable in the f.c.c. structure to at least 20 GPa, the highest hydrostatic pressure achieved. Pressure was limited to 20 GPa because of the large sample volume required to achieve both hydrostaticity and a reasonable diffraction signal from the low-atomic-number sample. Figure 2 also shows the least-squares fit to the Vinet⁹ EOS of the measured hydrostatic data. This results in an atmospheric-pressure bulk modulus $K_0 = 18.1 \pm 1.8\ \text{GPa}$ (compressibility $\kappa = (5.5 \pm 0.5) \times 10^{-2}\ \text{GPa}^{-1}$) and pressure derivative $dK_0/dP = 5.7 \pm 0.6$. Equally satisfactory fits are achieved with the Birch-Murnaghan EOS¹⁰ and yield the same values of K_0 and dK_0/dP . To 20 GPa, these semi-empirical equations have no difficulty fitting the C_{60} data, and we conclude that to this pressure there is no significant change in the bonding character of the fullerite. Down to an f.c.c. lattice parameter of $a = 12.45\ \text{\AA}$, we see no indication that the solid compressibility is being affected by the incompressibility of the C_{60} molecules. If the bonding between C_{60} molecules in the isotropic f.c.c. structure is identical to the inter-planar bonding in graphite⁵, we would expect a bulk modulus of 12

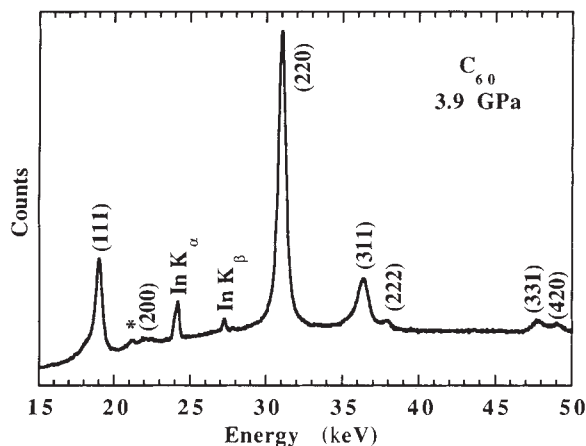


FIG. 1 Energy-dispersive X-ray diffraction spectrum of C_{60} fullerite at 3.9 GPa in experiment 3. This spectrum has been indexed to the f.c.c. lattice. The diffraction angle $\theta = 2.3904 \pm 0.0006^\circ$, and the peak marked with an asterisk is an escape peak from the (220) peak.

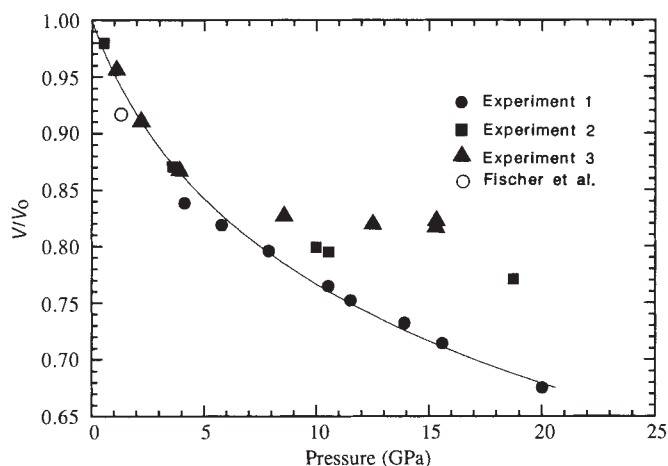


FIG. 2 The room-temperature equation of state of f.c.c. C_{60} fullerite as measured in the three experiments (see Table 1). The line represents a fit of the Vinet equation of state⁹ to the hydrostatic data of experiment 1. The single datum point of Fischer *et al.*¹² is also shown.

GPa. The larger bulk modulus measured for fullerite is consistent with the π -orbital rehybridization which results in a greater electron density outside the C_{60} molecule than inside⁴. Furthermore, the shorter bonds (at atmospheric pressure) between nearest-neighbour C_{60} molecules⁷ ($3.0\ \text{\AA}$) are expected to be stiffer than the inter-planar bonds in graphite¹¹ ($3.35\ \text{\AA}$).

At atmospheric pressure, the (200) diffraction peak is not observed in either powder or single-crystal X-ray diffraction patterns⁷. The reason for this can be easily seen by modelling the C_{60} molecules as hollow spheres of charge of radius r_c

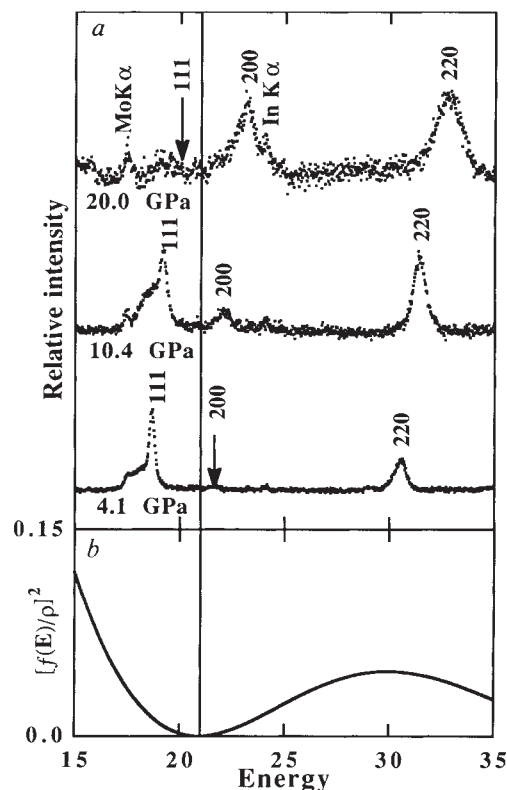


FIG. 3 a, Energy-dispersive X-ray diffraction spectra of C_{60} fullerite in experiment 1, showing the influence of pressure on the relative intensities of the f.c.c. reflections. The Mo and In fluorescences are from the gasket and detector, respectively. An inelastic scattering background has been numerically subtracted. b X-ray scattering form factor, equation (1), for hollow spheres of radius $r_c = 3.55\ \text{\AA}$ and ϵ equal to that used for the data in a.

centred on a f.c.c. lattice¹². The molecular form factor f for this radial charge density $U(r) = \rho\delta(r - r_c)$, where ρ is the uniform charge distribution on the sphere, can be written as¹³

$$f(E) = \int_0^\infty \rho\delta(r - r_c) \frac{\sin\left(\frac{2\pi r}{d}\right)}{\left(\frac{2\pi r}{d}\right)} dr = \frac{\rho\epsilon \sin\left(\frac{2\pi r_c E}{\epsilon}\right)}{2\pi r_c E} \quad (1)$$

where the constant $\epsilon \equiv Ed = hc/2 \sin \theta$, with E the diffraction energy and $d = a/\sqrt{h^2 + k^2 + l^2}$ the lattice spacing of the cubic-lattice hkl plane. Equation (1) has zeros at

$$\frac{a}{r_c} = \frac{2\sqrt{h^2 + k^2 + l^2}}{n} \quad (2)$$

with n an integer. For the atmospheric-pressure case⁷, where $a_0 = 14.198 \text{ \AA}$, the (200) peak will be absent for $r_c = 3.55 \text{ \AA}$, in close agreement with the experimentally⁷ determined value of $3.5 \text{ \AA}$ and the theoretical⁴ value of $3.55 \text{ \AA}$ for the free C_{60} molecule. If, however, the intramolecular bonds are significantly less compressible than the intermolecular bonding, then a/r_c will be pressure dependent, and the accidental absence of the (200) line will be eliminated at high pressure. This is shown in Fig. 3, where the (200) intensity increases as it shifts out of the region of $f = 0$. By 20 GPa, the (111) diffraction peak intensity has diminished nearly to zero. Although detailed analysis is complicated by powder orientation effects, these observed intensity changes imply a reduction in a/r_c of nearly 13% at 20 GPa. As a has diminished by more than 12% at this pressure, this implies no reduction in r_c within experimental error. This is consistent with the expected significantly lower compressibility of the C_{60} molecule relative to the bulk fullerite.

Figure 2 demonstrates that above about 5 GPa there are differences between hydrostatic and non-hydrostatic compression of solid C_{60} . These differences are not due to pressure gradients between the position of the ruby chip and the X-ray collimator position, as the pressure was measured at several places in the sample chamber. For example, a maximum gradient of 1 GPa over the entire sample was measured in experiment 2 at 18.7 GPa. In the geometry used here, the EDXD technique probes diffraction planes parallel to the stress axis of the diamond anvil cell. The existence of a higher stress in this direction will result in a larger measured lattice parameter and volume, determined by the Poisson ratio of fullerite. For this to explain the present data, the highly compressible fullerite must be capable of supporting uniaxial stresses. One possible mechanism for the required shear strength is the presence of elongated C_{70} molecules or other fullerenes in the lattice that act as dislocation pinning centres. Such an effect has been seen in graphite, where the basal-plane shear modulus c_{44} increases by nearly an order of magnitude when substitutionally doped with 1500 p.p.m. boron¹⁴. Further work, including diffraction-angle dependence of the measured lattice constant, is needed to confirm this explanation.

Both the C_{60} molecule and the f.c.c. fullerite structure are thus stable upon hydrostatic compression to at least 20 GPa. The non-hydrostatic experiments 2 and 3, however, show crystallographic transformations from the f.c.c. structure at $20 \pm 2 \text{ GPa}$ and $16.5 \pm 1.0 \text{ GPa}$, respectively. The difference in measured transition pressures is probably due to differences in the details of the stress state achieved in the two experiments. In experiment 3, where no pressure medium was used, the transition is characterized by an abrupt splitting of the (111) diffraction peak, broadening of the (311) peak, and no influence on the (220) peak to 24 GPa. Such behaviour indicates a transition to a structure of lower symmetry. The transition was observed well after the ruby spectra indicated the presence of a non-hydrostatic state of stress. The high-pressure phase cannot be indexed to a hexagonally close-packed unit cell, and further work is needed to characterize this phase and its reversibility.

Received 3 May; accepted 10 May 1991.

1. Krättschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. *Nature* **347**, 354–358 (1990).
2. Haddon, R. C. *et al.* *Nature* **350**, 320–322 (1991).
3. Hebard, A. F. *et al.* *Nature* **350**, 600–601 (1991).
4. Haddon, R. C., Brus, L. E. & Raghavachari, K. *Chem. Phys. Lett.* **131**, 165–169 (1986).
5. Hanfland, M., Beister, H. & Syassen, K. *Phys. Rev. B* **39**, 12589–12603 (1989).
6. Ruoff, R. S. & Ruoff, A. L. *Nature* **350**, 663–664 (1991).
7. Flemming, R. M. *et al.* *Proc. Mater. Res. Soc. Boston*, November, 1990 (Materials Research Society, Boston, 1991).
8. Mao, H. K., Bell, P. M., Shaner, J. W. & Steinberg, D. J. *J. appl. Phys.* **49**, 3276–3283 (1978).
9. Vinet, P., Ferrante, J., Smith, J. R. & Rose, J. H. *J. Phys. C* **19**, L467–L473 (1986).
10. Macdonald, J. R. & Powell, D. R. *J. Res. Natl Bur. Stand. A*, **75A**, 441–453 (1971).
11. Donohue, J. *The Structure of the Elements* (Wiley, New York, 1974).
12. Fischer, J. E. *et al.* *Science* (submitted).
13. James, R. W. *The Optical Principles of the Diffraction of X-rays* (Ox Bow, Woodbridge, Connecticut, 1982).
14. Soule, D. E. & Nezbeda, C. W. *J. appl. Phys.* **39**, 5122–5139 (1968).

ACKNOWLEDGEMENTS. This work is based upon data acquired at CHESS, supported by the NSF. We thank A. L. Ruoff for the use of his ruby system.

Simulations of the effect of a warmer climate on atmospheric humidity

Anthony D. Del Genio, Andrew A. Lacis & Reto A. Ruedy

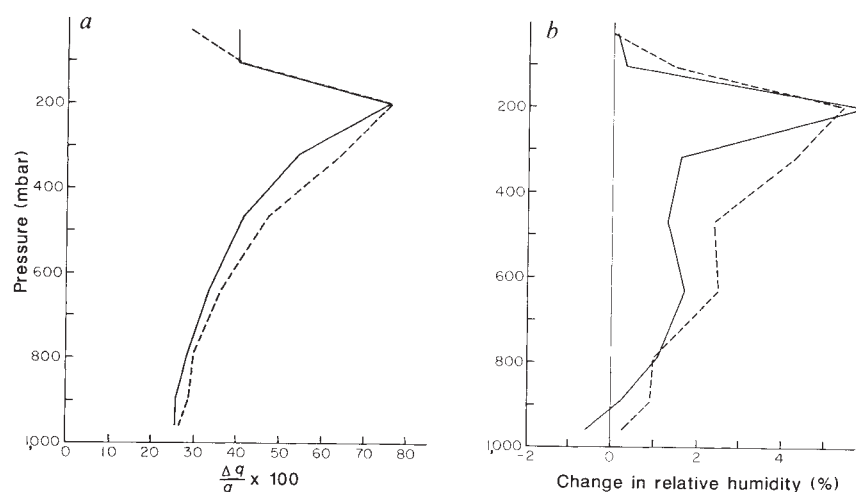
NASA/Goddard Institute for Space Studies, New York, New York 10025, USA

THE increases in the concentration of water vapour constitute the single largest positive feedback in models of global climate warming caused by greenhouse gases^{1,2}. It has been suggested^{3–5} that sinking air in the regions surrounding deep cumulus clouds will dry the upper troposphere and eliminate or reverse the direction of water vapour feedback. We have now tested this hypothesis by performing an idealized simulation of climate change with two different versions of a climate model. The versions differ in their parameterizations of moist convection and stratiform clouds, but both incorporate the drying due to subsidence of clear air. Despite increased drying of the upper troposphere by cumulus clouds, upper-level humidity increases in the warmer climate because of enhanced upward moisture transport by the general circulation and increased accumulation of water vapour and ice at cumulus cloud tops. The model behaviour is consistent with recent satellite estimates of the water vapour feedback^{6,7}.

We express water vapour concentration in terms of the specific humidity q , defined as the ratio of the mass density of water vapour to that of moist air. $q = q_s r$, where q_s , the saturation humidity, is the maximum possible humidity in equilibrium over a flat surface of water, and r , the relative humidity, is the fraction of the saturation humidity actually realized. q_s is a function only of temperature and pressure and is determined by the Clausius–Clapeyron equation of thermodynamics⁸; r is the net result of competing dynamic, radiative and microphysical processes. The values of q_s increase nonlinearly with temperature. As a result, the q of near-surface tropical air is two to three orders of magnitude greater than that of the colder upper troposphere and polar regions.

Evaporation from the surface and precipitation are the ultimate source and sink of atmospheric moisture, respectively. Within the atmosphere, several processes redistribute water vapour vertically and horizontally. Deep precipitating cumulus-updrafts lift moisture from near the surface and inject some of it into the upper troposphere as vapour or ice, a process known as detrainment. The clear environment surrounding the updrafts slowly sinks in response; this compensating subsidence dries the atmosphere by bringing low- q air down from above. Condensation in clouds is a local sink of water vapour, and re-evaporation of falling precipitation is a local source. Large-scale

FIG. 1 *a*, Vertical profiles of the global mean fractional change in specific humidity q for the last month of the GCM simulations; the difference is computed as the warmer (+2 K) minus the cooler (−2 K) perturbed SST climate. *b*, The corresponding profiles of the absolute change in relative humidity, obtained by averaging instantaneous values of r over the month. The associated mean upper-troposphere temperature change is 5.3 and 5.6 K in the old and new versions of the GCM, respectively. Note that for water vapour concentration in the upper troposphere to decrease as climate warms, as predicted by several authors^{3–5}, relative humidity would have to decrease by more than 20%. Continuous trace, model II; dashed trace, new convection and clouds.



eddies transfer water vapour upwards from low altitudes, especially in mid-latitude storms where moist air rises near warm fronts and drier air sinks behind cold fronts. The same is true of the mean meridional circulation (MMC), dominated by the Hadley cell, whose rising branch is moist and sinking branch drier.

Our two simulations used the general circulation model (GCM) of the NASA/Goddard Institute for Space Studies. One used the standard version of the GCM, called model II⁹. Model II contains a penetrative cumulus scheme with representations of compensating subsidence, detrainment and evaporation of falling precipitation; the cumulus mass flux in unstable conditions is 50% of the mass of the cloud base layer. This differs from traditional convective adjustment¹⁰ and cloud–environment mixing (Kuo-type)¹¹ schemes, which directly mix moisture upward. Stratiform clouds are diagnosed by assigning a temperature variance on scales smaller than that of the grid and computing the supersaturated fraction within each grid box.

The second simulation used a new version of the GCM with improved representations of cumulus^{12,13} and stratiform cloud¹⁴ based on recent tropical field studies and validated against a variety of satellite data sets; in other respects, this model is identical to model II. The new cumulus model includes computation of mass flux designed to produce a quasi-equilibrium¹⁵ between convective-scale and large-scale motions, simultaneous deep and shallow convection, and transport by cumulus-scale

downrafts as well as environmental subsidence. Calculations for stratiform clouds are based on a cloud liquid–ice water budget, including a representation of mesoscale cumulus anvils, different microphysical properties for liquid and ice, collection of cloud water by precipitation, diffusional growth of ice, cloud-top entrainment instability and variable optical thickness. Evaporation of condensate is largely restricted to the lower troposphere, beneath the bases of cumulus anvils, and this allows a more direct test of the effects of cumulus subsidence in the upper troposphere. Both models were run with $8^\circ \times 10^\circ$ horizontal resolution and nine vertical levels, including three in the upper troposphere (layers centred at 201, 320 and 468 mbar).

The idealized simulation of climate change was identical to that conducted in a recent comparison of 19 GCMs¹⁶. Both models were perturbed with globally uniform −2 K and +2 K increments of sea surface temperature (SST), and the differences between the warmer and cooler climates compared. The models were run for one year as if conditions for July persisted throughout. Quantities of sea ice were fixed. Although this differs in several ways from GCM simulations of increasing trace gases, the changes in moisture budget that we describe are consistent with those in an experiment with model II that considered doubled carbon dioxide concentrations².

Figure 1 shows the changes in specific and relative humidity profiles (warmer minus cooler) that occur in the two versions of the GCM. The concentration of water vapour increases substantially at all altitudes as the climate warms. The greatest fractional enhancement occurs in the upper troposphere, although the greatest absolute increase is near the surface. Relative humidity is almost constant in the lower troposphere and increases by a few per cent at higher levels. The new version of the GCM has slightly stronger water vapour feedback than Model II: a feedback analysis using a one-dimensional radiative–convective model² indicates that the humidity increases in Fig. 1 contribute 2.69 K to the warming in model II and 2.98 K in the new version of the GCM. The changes predicted by the two models are nonetheless strikingly similar, given the substantial differences in the physics employed. The small changes in r in both simulations imply that climate-induced changes in q are largely due to changes in q_s . This suggests that water vapour feedback is controlled primarily by temperature and is not very sensitive to the fine-tuning of the cloud models. Six other GCMs with mass-flux convection schemes also give strong positive water vapour feedback when subjected to the same SST perturbation¹⁶.

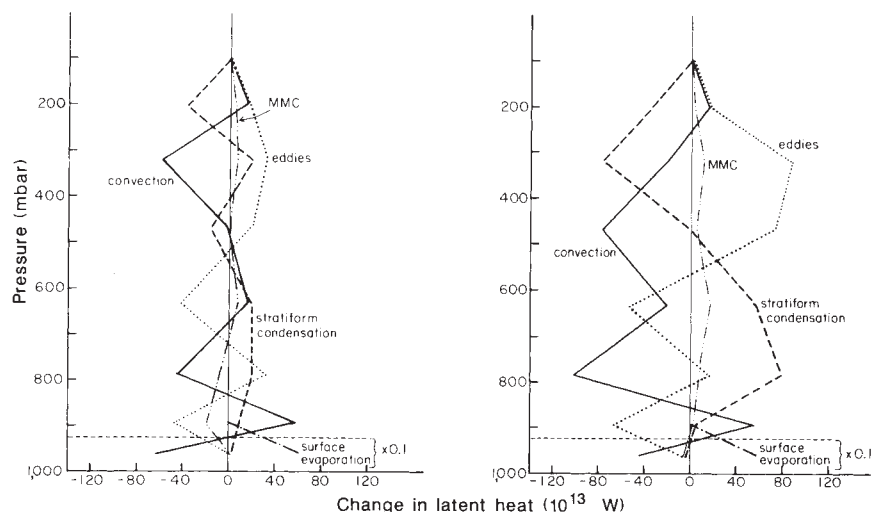
The reasons for positive water vapour feedback are made clear by examining changes in the individual terms of the atmospheric moisture budget between the warmer and cooler climate (Fig. 2). In equilibrium, there is no net moisture storage, so the

TABLE 1 Global diagnostics of general circulation strength and specific humidity gradient

	Model II		New convection and clouds	
	−2 K	+2 K	−2 K	+2 K
Cumulus mass flux (10^{11} kg s ^{−1})	18.8	17.8	11.7	11.3
Eddy kinetic energy (10^4 J m ^{−2})	91.5	86.7	103.3	107.1
Peak Hadley cell streamfunction (10^9 kg s ^{−1})	184	172	156	160
Specific humidity difference 468 mbar – 201 mbar (10^{-4} kg per kg)	11.4	16.9	11.0	17.5
4°N – 59°N (10^{-4} kg per kg)	14.0	24.2	16.1	23.0

The numbers are for the final month of the simulations for both versions of the GCM. Convective moisture transport depends only on the vertical humidity gradient; transport by large-scale motions is affected by both horizontal and vertical gradients.

FIG. 2 Globally averaged vertical profiles of the changes (warmer minus cooler) in the individual components of the moisture budget (expressed in units of change of latent heat per unit time) for the last month of the GCM simulations for model II (left) and the new version of the GCM (right). The abscissa represents the climate change caused by each process in the time derivative of water vapour mass multiplied by the latent heat of condensation of water. The full lines represent both moist (cumulus clouds) and dry (primarily boundary layer turbulence) convection, almost exclusively moist above 600 mbars. The dotted lines include transient and stationary eddies, the tropical Walker circulation and the GCM representation of mesoscale cumulus updrafts and downdrafts. The MMC (dot-dashed lines) is dominated by the Hadley cell, with small contributions from the Ferrel and polar cells. The dashed lines represent stratiform condensation, including both sinks (cloud formation) and sources (evaporation of falling precipitation) of vapour resulting from phase changes. At the lowest model level, the abscissa has been multiplied by 0.1 so that, for example, a term in surface evaporation which is $\sim 600 \times 10^{13} \text{ W}$ is plotted at $\sim 60 \times 10^{13} \text{ W}$.



net effect of the individual processes shown in Fig. 2 sum to zero at each level. A positive (negative) change in the figure means that the process in question moistens (dries) the warmer climate more than the cooler climate. In both versions of the GCM, cumulus-induced subsidence increases drying in the warmer climate at 320 and 468 mbar, as anticipated by Ellsaesser³ and Lindzen^{4,5}, but this is offset by increased moistening as a result of net upwards transport by large-scale eddies and the MMC. The eddies and MMC oppose convection because they converge moisture throughout the upper troposphere, whereas convection does so only at cloud top. A high proportion (roughly two-thirds according to model II, one-half according to the new version of the GCM) of the increased global eddy moistening at 320 mbar occurs outside the latitudes of tropical convection, much of it in mid-latitude storms. Global budgets therefore cannot be reliably inferred by considering tropical processes in isolation. At the tropopause, convection itself increasingly moistens the warmer atmosphere, because detrainment of saturated cloud air and ice occurs preferentially at higher altitude, and with increased condensate, as the climate warms.

Upper-troposphere humidity is maintained by transport of water down its gradient from lower to upper troposphere. Thus, upper-level water vapour increases in the GCM only if transports by the individual processes increase in the warmer climate. Dynamical moisture transport depends on the product of a mass flux (density times velocity) and the specific humidity gradient. Table 1 shows that the mass fluxes associated with various components of the general circulation change little, and sometimes decrease, as climate warms. The increased moisture transport is thus caused solely by marked increases in the vertical and horizontal humidity gradients, which are tightly coupled to temperature. In other words, upper-troposphere humidity increases in a warmer climate, because a general circulation that is almost unchanged will transport water upwards along a much steeper gradient.

Recent satellite data sets confirm the GCM view of water vapour feedback on climate. The SAGE II solar occultation instrument, for example, indicates that water vapour concentration is higher in summer than winter in both hemispheres at all tropospheric altitudes and latitudes, regardless of whether the comparison is made at a location of rising or sinking in the MMC⁶. Seasonal differences in relative humidity measured by SAGE II are largest in the upper troposphere (typically 5–10% higher in summer), consistent with the predictions of the GCM. Longwave radiation fluxes measured in clear-sky regions by the

Earth Radiation Budget Experiment imply that the contribution of water vapour to the greenhouse effect is largest in the warmest ocean regions, regardless of its vertical distribution⁷. Finally, radiosonde moisture profiles indicate higher amounts of water vapour both in summer and in the warm phase of the El Niño cycle in the middle and upper troposphere^{17,18}, although the uncertainty in specific humidity at these altitudes is considerable. In summary, the available observations indicate higher upper-troposphere humidity in a warmer atmosphere, the opposite result to that expected from consideration of cumulus subsidence alone.

Our findings have several implications for climate monitoring strategies and global warming policy. As water vapour produces the largest known positive feedback on climate, it is essential to refine our estimates of its impact by making more accurate observations of upper-troposphere water vapour and radiative fluxes. The large increases in humidity seen in Fig. 1 also indicate that long-term monitoring of water vapour by satellites may be a useful means of detecting a global warming trend¹⁹. The general agreement between GCM and observational estimates of water vapour feedback also gives us an estimate of the minimum climate sensitivity to human activity. Specifically, an equivalent doubling of CO₂ can be expected to produce an equilibrium global warming of at least 1.5–2.0 °C (using the most negative estimate of cloud feedback produced by a GCM²⁰), even more if cloud and snow or ice feedbacks are nearly neutral or positive. □

Received 28 January; accepted 11 April 1991.

1. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **24**, 241–259 (1967).
2. Hansen, J. et al. *Climate Processes and Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) 130–163 (American Geophysical Union, Washington DC, 1984).
3. Ellsaesser, H. W. *Atmos. Envir.* **18**, 431–434 (1984).
4. Lindzen, R. S. *Bull. Am. met. Soc.* **71**, 288–299 (1990).
5. Lindzen, R. S. *Bull. Am. met. Soc.* **71**, 1465–1467 (1990).
6. Rind, D. et al. *Nature* **349**, 500–503 (1991).
7. Raval, A. & Ramanathan, V. *Nature* **342**, 758–761 (1989).
8. Wallace, J. M. & Hobbs, P. V. *Atmospheric Science—An Introductory Survey* (Academic, New York, 1977).
9. Hansen, J. et al. *Mon. Weath. Rev.* **111**, 609–662 (1983).
10. Manabe, S. et al. *Mon. Weath. Rev.* **93**, 769–798 (1965).
11. Kuo, H. L. *J. Atmos. Sci.* **31**, 1232–1240 (1974).
12. Del Genio, A. D. & Yao, M.-S. *J. Atmos. Sci.* **45**, 2641–2668 (1988).
13. Yao, M.-S. & Del Genio, A. D. *J. Clim.* **2**, 850–863 (1989).
14. Del Genio, A. D. & Yao, M.-S. *AMS Conf. Cloud Phys. Preprint No. 9A.4*, 497–504 (American Meteorological Society, Boston, 1990).
15. Arakawa, A. & Chen, J.-M. *Short- and Medium-Range Numerical Weather Prediction*, Suppl. to *J. met. Soc. Japan*, 107–131 (1987).
16. Cess, R. D. et al. *J. geophys. Res.* **95**, 16601–16615 (1990).
17. Oort, A. H. *NOAA Prof. Paper No. 14* (US Dept. Commerce, Washington DC, 1983).
18. Pan, Y.-H. & Oort, A. H. *Mon. Weath. Rev.* **111**, 1244–1258 (1983).

19. Hansen, J., Rossow, W. & Fung, I. *Issues Sci. Tech.* **7**, 62–69 (1990).
 20. Mitchell, J. F. B. *et al. Nature* **341**, 132–134 (1989).

ACKNOWLEDGEMENTS. We thank R. Cess, J. Hansen, R. Miller, W. Rossow, D. Starr, P. Stone and M.-S. Yao for helpful comments. This work was supported by the DOE Interagency Agreements under the Quantitative Links and Atmospheric Radiation Measurement Programs. R.A.R. is supported by ST Systems Corporation through a support contract with NASA.

Enhanced particle fluxes in Bay of Bengal induced by injection of fresh water

V. Ittekkot*, R. R. Nair†, S. Honjo‡, V. Ramaswamy†, M. Bartsch*, S. Manganini‡ & B. N. Desai†

* Institute of Biogeochemistry and Marine Chemistry, University of Hamburg, Bundesstrasse 55, 2000 Hamburg 13, Germany
 † National Institute of Oceanography, Dona Paula, Goa 403004, India
 ‡ Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

THE melting of ice sheets during deglaciation results in the injection of large amounts of fresh water into the oceans¹. To investigate how such injections might influence particle fluxes in the ocean, and hence the uptake of atmospheric CO₂, we deployed three sediment-trap moorings (two traps in each mooring) in the northern, central and southern parts of the Bay of Bengal, respectively. The Bay of Bengal is suitable for such a study, because some of the world's largest rivers² supply pulses of fresh water and sediment to the bay, resulting in large seasonal changes in surface salinity³. We find that the maximum river discharge, which occurs during the southwest monsoon, coincides with the maximum observed flux of particulate matter. From north to south, the carbonate flux increases, whereas fluxes of opal, organic carbon and particulate matter decrease. The overall flux pattern seems to be controlled by the seasonally varying input from the rivers and the accompanying shift in marine biogenic production. We conclude that freshwater pulses during deglaciation may therefore have caused similar

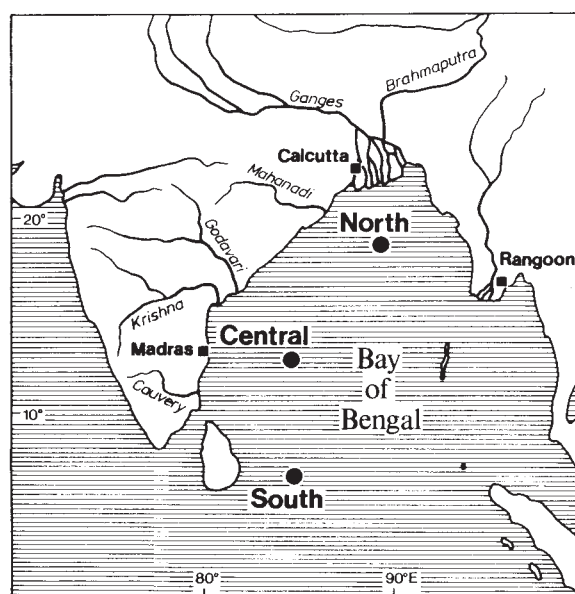


FIG. 1 Locations of sediment-trap moorings in the northern (17°26'N, 89°35'E; water depth 2,263 m, trap depths 809 and 1,727 m), central (13°09'N, 84°21'E; water depth 3,259 m, trap depths 906 and 2,282 m) and southern (04°26'N, 87°19'E; water depth 4,017 m, trap depths 1,040 and 3,006 m) Bay of Bengal. Each mooring consists of two time-series sediment traps¹⁷. They were first deployed in October 1987 and have been recovered and redeployed three times since October 1987 using the research vessels RV *Sonne* and ORV *Sagar Kanya*.

shifts in marine biogenic production, resulting in short-term episodes of increased oceanic uptake of atmospheric CO₂.

Surface salinity in the Bay of Bengal fluctuates over a wide range of values, both seasonally and geographically⁴. The greatest change occurs towards the end of the southwest monsoon, when water discharge from the Ganges and Brahmaputra and from the peninsular Indian rivers is at a maximum. During this period (August–November) mean salinity decreases by more than 7‰ (ref. 4).

The results from the first year of our continuing experiment in the Bay of Bengal show seasonal fluctuations in particle fluxes at all three sites (Figs 1 and 2). Forty to fifty per cent of the total annual flux occurs during the southwest monsoon, which covers 3–4 months. During the rest of the year (northeast monsoon and intermonsoon periods), particle fluxes are uniformly low. An exception is one sampling interval during

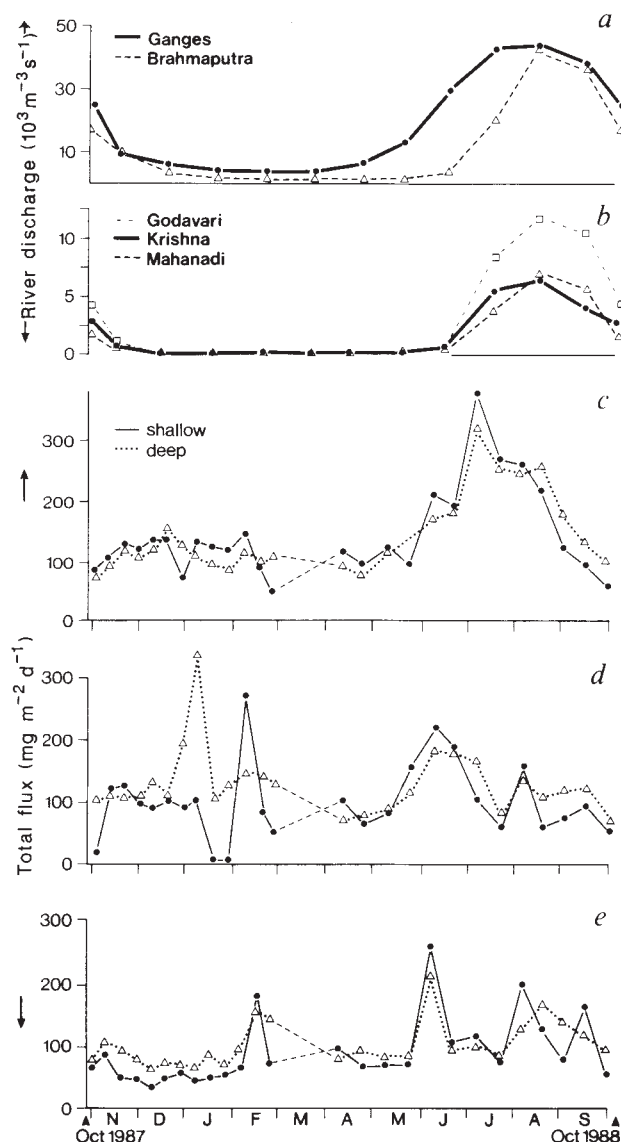


FIG. 2 Time-series data of particle flux in the Bay of Bengal for October 1987 to October 1988. *a, b*, River discharge data from ref. 18. *c, d, e*, Total particle flux from north, central and south Bay of Bengal, respectively. The sediment traps were programmed to measure the flux of sinking particles at intervals of 9.5 days during the first deployment and 14.5 days during the second deployment. The collecting cups were poisoned with mercury chloride before deployment. On recovery of the traps, the samples were wet-sieved and split using a precision rotary splitter. One-quarter of the <1-mm size fraction was filtered through preweighed Nuclepore filters (0.5 µm) and dried at 40 °C. This fraction was used for total flux calculations.

TABLE 1 Fluxes of components to the deep Bay of Bengal and their ratios

Period	Trap	C _{carb} flux		Opal flux		C _{org} flux		Total nitrogen flux		Lithogen flux		Total flux	C/N	C _{org} /C _{carb}	Carb/Opal
		(g m ⁻²)	(%)	(g m ⁻²)	(%)	(g m ⁻²)	(%)	(g m ⁻²)	(%)	(g m ⁻²)	(%)				
North															
SW-NE	shallow	1.44	26.4	0.75	13.7	0.39	7.1	0.05	0.9	2.54	46.6	5.44	7.76	2.40	1.88
	deep	1.18	19.6	0.96	16.0	0.29	4.8	0.04	0.6	3.32	55.3	6.00	8.17	2.18	1.32
NE	shallow	3.39	34.5	1.76	17.9	0.69	7.0	0.08	0.8	3.40	34.6	9.80	8.23	2.27	2.05
	deep	2.54	25.7	1.40	14.2	0.53	5.4	0.06	0.6	4.93	50.1	9.85	8.27	1.91	1.94
NE-SW	shallow	2.39	27.4	1.60	18.3	0.73	8.4	0.07	0.8	3.39	38.8	8.72	-	2.55	1.86
	deep	2.49	27.5	1.53	16.9	0.53	5.8	0.06	0.6	4.06	44.8	9.06	7.95	1.99	1.67
SW	shallow	4.33	16.8	5.70	22.1	1.78	6.8	0.23	0.8	12.56	48.6	25.81	7.71	3.52	1.04
	deep	4.51	16.9	4.17	15.6	1.30	4.8	0.15	0.5	15.65	58.6	26.68	8.68	2.39	1.27
Total annual	shallow	11.55		9.81		3.59		0.43		21.89		49.77			
	deep	10.72		8.06		2.65		0.31		27.96		51.59			
Central															
SW-NE	shallow	1.99	44.3	0.58	12.9	0.32	7.1	0.04	0.8	1.33	29.6	4.49	7.03	1.90	2.97
	deep	2.43	41.5	0.84	14.3	0.29	4.9	0.03	0.5	2.04	34.8	5.85	8.72	1.10	2.82
NE	shallow	4.06	50.7	1.51	18.8	0.51	6.3	0.06	0.7	1.49	18.6	8.00	7.55	1.36	2.57
	deep	5.39	38.0	3.74	26.3	0.88	6.0	0.11	0.7	3.46	24.1	14.18	7.95	1.41	1.98
NE-SW	shallow	3.40	39.6	1.82	21.2	0.75	8.8	0.09	1.0	1.98	23.1	8.57	8.37	2.05	1.90
	deep	3.36	39.7	1.47	17.4	0.53	6.3	0.06	0.7	2.66	31.4	8.46	8.72	1.47	2.67
SW	shallow	5.18	36.7	2.43	17.2	1.06	7.5	0.13	0.9	4.57	32.4	14.09	8.28	1.72	2.23
	deep	5.74	35.2	2.37	14.5	0.91	5.5	0.10	0.6	6.54	40.1	16.30	8.76	1.35	2.48
Total annual	shallow	14.63		6.34		2.64		0.31		9.37		35.15			
	deep	16.92		8.42		2.61		0.30		14.70		44.79			
South															
SW-NE	shallow	1.89	51.3	0.34	9.2	0.18	4.0	0.02	0.5	1.11	30.1	3.68	7.62	0.83	5.67
	deep	2.57	47.4	1.15	21.2	0.24	4.4	0.03	0.5	1.25	23.1	5.42	7.91	1.05	2.19
NE	shallow	2.20	39.0	0.65	11.5	0.34	6.0	0.04	0.7	2.16	38.2	5.64	7.51	1.26	3.83
	deep	3.47	42.8	1.79	22.9	0.39	4.8	0.04	0.5	2.13	26.2	8.10	7.78	0.91	1.93
NE-SW	shallow	3.36	45.9	1.36	18.5	0.57	7.7	0.07	0.9	1.55	21.2	7.32	7.80	1.40	3.71
	deep	3.94	43.8	1.43	15.9	0.54	6.0	0.06	0.6	2.62	29.1	8.98	8.12	1.08	2.85
SW	shallow	8.63	53.3	1.72	10.6	1.20	7.4	0.15	0.9	3.67	22.4	16.19	7.50	1.09	5.37
	deep	8.52	54.6	2.90	18.6	0.87	5.5	0.10	0.6	2.56	16.4	15.58	8.18	0.87	2.91
Total annual	shallow	16.08		4.07		2.37		0.23		8.49		32.83			
	deep	18.50		7.27		2.04		0.28		8.56		38.08			

Data are for the fraction of size <1 mm and are given for the southwest (June–September) and northeast (November–February) monsoons, and for the period between monsoons (see also legend to Fig. 3).

February in the central and southern stations during which higher fluxes are recorded. The total fluxes, and the variability of these fluxes within sites, decreased from north to south. The flux maximum during the southwest monsoon consists of two peaks: the first, at the beginning of June, is related to meltwater influx, and the second is related to increased input of riverine sediment. During the southwest monsoon, nutrient inputs from rivers, and from entrainment of subhalocline layers through turbulent mixing, will increase primary production, especially in river plumes. According to previous studies^{5,6} increased primary production within river plumes may enhance particle flux to deep, offshore waters. Furthermore, if river-derived mineral particles are incorporated in large organic aggregates derived from primary production, this may accelerate particle sedimentation in the sea^{7,8}.

Component fluxes for opal, lithogenics and organic carbon decreased from north to south, and the carbonate flux increased (Fig. 3). Opal fluxes are higher than those reported from similar latitudes and depths⁹ except areas in the Arabian Sea directly affected by coastal upwelling¹⁰. On the other hand, organic carbon (C_{org}) fluxes to the deep Bay of Bengal at all sites are consistently higher than that reported from the Arabian Sea, an area conventionally considered¹⁰ to be one of the most productive regions of the world. Biogenic matter (carbonate, opal and organic matter) made up 42–56% of the total fluxes in the northern, and 67–77% in the central and southern stations. The ratio of C_{org}/C_{carb} , where C_{carb} is carbonate carbon, varied from 2–5 in the north to <1 in the south during the southwest monsoon. It has been suggested¹¹ that C_{org}/C_{carb} ratios of settling particles ('rain ratios') are a function of fertility, whereby ratios of ~0.6 are typical of deep-water particulate fluxes in low fertility areas and ratios of ~2 are typical of shallow traps. The ratio of carbonate to opal increases from north to south (Table 1) during all seasons and at all depths. It essentially reflects the prevalent plankton community structure in the upper ocean. In areas where high nutrient concentrations are available, for example in the northern Bay of Bengal, this ratio is shifted in

favour of opal^{12,13}. Comparison of these two ratios suggests that high C_{org} fluxes to the deep Bay of Bengal could result from both marine and riverine sources, where the marine source is essentially derived from diatoms, which indeed are the dominant phytoplankton species in these samples.

Within each site and depth the seasonally averaged flux pattern differs (Table 1). The decrease in flux between shallow and deep traps in the north results from a reduction in opal between the two. This may be due either to dissolution or to its lateral removal. In contrast, at the southern station opal fluxes increase with depth during the southwest monsoon, suggesting a lateral input of biogenic opal in mid-water. The absolute flux of lithogenic matter increases with depth for all seasons at all sites.

Time-series data (Fig. 2) show, however, that particle flux to the deep trap is generally lower than that to the shallow trap during peak fluxes, except for two sampling intervals during the northeast monsoon at the central and southern stations. This implies loss, dissolution or decomposition between traps. During low-flux periods, the deep traps recorded slightly higher values, resulting in an overall increase in fluxes averaged for seasons to deep traps (see above and Table 1). The greatest increases are for lithogenic fluxes at the northern and central stations, and result from scavenging from the water column of lithogenic matter by settling particle aggregates^{7,14,15}. We saw no evidence in the northern and central Bay of Bengal for previously suggested mechanisms such as chemolithotrophic production in midwater and zooplankton activity¹⁶ which should have resulted in an increase in organic matter fluxes to the deep traps. Zooplankton activity, however, could explain the increase in opal fluxes during the southwest monsoon in the southern station.

Our results from a marine region with seasonally changing surface salinity caused by freshwater discharge from rivers show large variations in both the quantity and the nature of particle fluxes to the deep ocean. The quantity of the fluxes is largely controlled by nonbiogenic matter derived from the land. Variations in the biogenic flux pattern show seasonal shifts from carbonate to opal domination. Carbon cycling in the ocean is

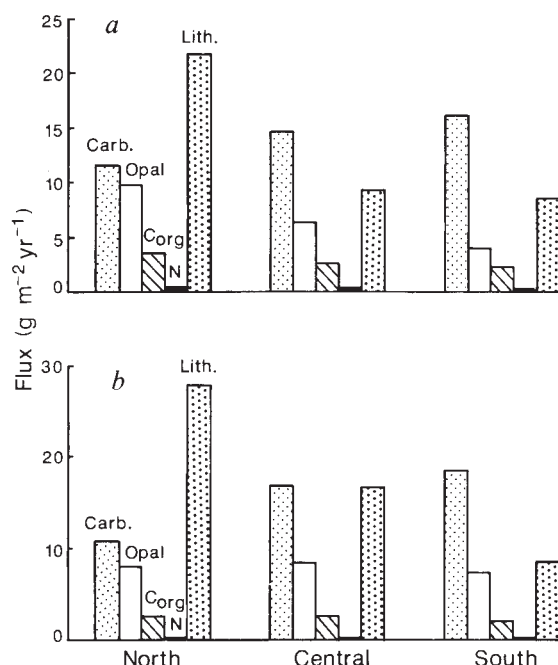


FIG. 3 Component fluxes to *a*, shallow and *b*, deep traps in the Bay of Bengal. Components were analysed for carbonate, opal, organic carbon and nitrogen by methods described elsewhere^{19,20}. Lithogenic fluxes were calculated as the difference between the total fluxes and the sum of carbonate, opal and organic matter ($C_{org} \times 1.8$) fluxes.

determined in part by the composition of primary producers: the settling of particulate organic carbon decreases the total carbon and partial pressure of CO_2 of the surface layers, whereas the removal of carbonate from surface waters aids in increasing atmospheric CO_2 by shifting the carbonate equilibria^{11,13}. Thus, differences in the initial production ratio of calcium carbonate to organic carbon can effect significant variations in atmospheric CO_2 . The shifts in upper-ocean plankton community structure, with our study indicating a dominance of diatoms over nanoplankton (low carbonate/opal ratios), have led to high C_{org}/C_{carb} ratios (up to 5 during the southwest monsoon when up to 50% of the annual flux occurs), which resemble values usually found in shallow traps^{11,13}. Previous model calculations¹³ indicate that a decrease of ~ 12 p.p.m. in atmospheric CO_2 can be effected just by doubling the C_{org}/C_{carb} ratio globally. The Bay of Bengal ratios associated with individual sampling intervals during the southwest monsoon are twice those measured at other times of the year. The changes are especially large in the north, which shows very marked salinity changes⁴.

Our results therefore have implications for research on the recent history of the Earth's climate, specifically that relating to short-term CO_2 fluctuations in the atmosphere during glacial-interglacial transitions. For example, the retreat of the ice sheets during the last deglaciation proceeded stepwise, with at least two large influxes of fresh water into the oceans accompanied by changes in the pattern of ocean circulation¹. In analogy to our above-mentioned results for a marine region, such fresh-water inputs might have caused short-term changes in marine-carbon removal processes of climatic significance. □

Received 22 November 1990; accepted 18 April 1991.

1. Fairbanks, R. G. *Nature* **342**, 637–642 (1989).
2. Milliman, J. D. & Meade, R. H. *Geology* **91**, 1–21 (1983).
3. Wyrtki, K. *Oceanographic Atlas of the International Indian Ocean Expedition* (National Science Foundation, Washington DC, 1971).
4. LaViolette, P. E. *Temperature, Salinity and Density of the World's Seas: Bay of Bengal and Andaman Sea*, Informal Rep. No. 67-57 (Naval Oceanographic Office, Washington DC 1967).
5. Deuser, W. G., Muller-Karger, F. E. & Hemleben, C. *J. geophys. Res.* **93**, 6857–6862 (1988).
6. Muller-Karger, F. E., McLain, C. R. & Richardson, P. L. *Nature* **333**, 56–59 (1988).

7. Ramaswamy, V., Nair, R. R., Manganini, S., Haake, B. & Ittekkot, V. *Deep-Sea Res.* **38**, 169–184 (1991).
8. Haake, B. & Ittekkot, V. *Naturwissenschaften* **77**, 75–79 (1990).
9. Honjo, S. in *Polar Oceanography* (ed. Smith, W. O. Jr) 687–739 (Academic, New York, 1990).
10. Nair, R. R. et al. *Nature* **338**, 749–751 (1989).
11. Berger, W. H. & Keir, R. S. in *The Carbon Cycle and Atmospheric CO_2 : Natural Variations Archaen to Present* (eds Sundquist, E. & Broecker, W. S.) 337–351 (American Geophysical Union, Washington, DC, 1985).
12. Ittekkot, V. & Zhang, S. *Global biogeochem. Cycles* **3**, 383–391 (1989).
13. Dymond, J. & Lyle, M. *Limnol. Oceanogr.* **30**, 699–712 (1985).
14. Honjo, 1982 *Science* **218**, 883–885 (1982).
15. Deuser, W. G., Brewer, P. G., Jickells, T. D. & Commeau, R. D. *Science* **214**, 388–391 (1983).
16. Walsh, I., Dymond, J. & Collier, R. *Deep-Sea Res.* **35**, 43–58 (1988).
17. Honjo, S. & Doherty, K. W. *Deep-Sea Res.* **35**, 133–149 (1988).
18. *Discharge of Selected Rivers of the World* (UNESCO, Paris, 1979).
19. Honjo, S., Manganini, S. J. & Cole, J. J. *Deep-Sea Res.* **29**, 609–625 (1982).
20. Michaelis, W. & Ittekkot, V. in *Transport of Carbon and Minerals in World Rivers Part 1* (ed. Degens, E. T.) 233–243 (Mitt. Geol. Palontol. Inst. Univ. Hamburg, SCOPE/UNEP Sonderbd. 52, 1982).

ACKNOWLEDGEMENTS. We thank V. Rajaraman, D. Gracias, F. Vijayan and P. Jöhndt for technical assistance and the officers and crew of the research vessels RV *Sonne* and ORV *Sagar Kanya* for help in deployment and recovery of traps. Financial support was provided by the Federal German Ministry for Research and Technology, Bonn, and the Council of Scientific and Industrial Research, New Delhi, and ship time on ORV *Sagar Kanya* by the Department of Ocean Development (New Delhi).

Eruptive and diffuse emissions of CO_2 from Mount Etna

P. Allard*, J. Carboneille†, D. Dajlevic†, J. Le Bronec†, P. Morel†, M. C. Robe†, J. M. Maurenas‡, R. Faivre-Pierret‡, D. Martin§, J. C. Sabroux* & P. Zettwoog†

* Centre des Faibles Radioactivités, CNRS-CEA, 91198 Gif-sur-Yvette, France

† Commissariat à l'Energie Atomique, IPSN, SPIN/DPT, 91198 Gif-sur-Yvette, France

‡ Commissariat à l'Energie Atomique, LESI, 38041 Grenoble Cedex, France

§ Centre de Recherches en Physique de l'Atmosphère, EERM, 78470 Magny-les-Hameaux, France

MOUNT Etna, in Sicily, is one of the world's most actively degassing volcanoes¹. Here we use data collected from 1975 to 1987 to estimate carbon dioxide emissions from the summit craters and the upper flanks of the volcano. By combining measurements of the SO_2 flux in the plume (refs 1–6 and this paper) with measurements of the CO_2/SO_2 ratio of the plume gases, we find that the average output of CO_2 from summit crater degassing is 13 ± 3 Tg yr⁻¹. This is an order of magnitude higher than the annual CO_2 output from Kilauea^{7,8}, Hawaii, and representative arc volcanoes^{9,10}. Furthermore, we find that diffuse emissions of CO_2 from the upper flanks of Etna are magma-derived and are of a similar magnitude to those emitted from the crater plume. This observation, as well as others^{11–14}, verifies the idea¹⁵ that extensive diffuse release of magmatic CO_2 may occur in volcanically active regions—a process that needs to be taken into account when evaluating the volatile budget of subaerial volcanism. Such degassing may be of use for monitoring volcanic activity, could provide a means for radiocarbon dating of eruptions, and may be a mechanism by which CO_2 is injected into crater lakes.

Mount Etna is a large strato-volcano (Fig. 1) producing alkaline basaltic lava during summit and flank eruptions¹⁶. A prominent feature of its activity is the persistence of voluminous summit crater degassing, which occurs during and between eruptions. Our analyses of both the high-temperature crater gases (refs 17, 18 and P.A., manuscript in preparation) and the dilute plume^{2,4,5,19} show this exhalation to consist mainly of water vapour, CO_2 , SO_2 and halogen compounds. The high CO_2/SO_2 molar ratio (≥ 10) in the exhalation probably reflects both a high abundance of CO_2 in the alkaline magma and a deep exsolution²⁰ of CO_2 from it.

Between 1975 and 1985 we performed repeated measurements of the SO_2 plume flux, using remote correlation spectrometry from a ground vehicle and a laboratory aircraft^{1–6}. Our results

TABLE 1 SO₂ and CO₂ fluxes (10⁹ g d⁻¹) in Mount Etna summit crater plume

Date	SO ₂				CO ₂			Activity level [§]	Reference
	range	<i>n</i>	mean	s.d.	<i>n</i>	mean	s.d.		
1975 (06/14–30)	2.7–4.8	17	3.8	0.42				M	1
1976 (01/24–02/11)	5.4–12.4	3	8.4	3.60				H	1
(06/13–27)	1.3–6.5	8	3.3	1.95	1	86 (23)*	—	L to M	1, 2
1977 (05/22–23)	1.1–1.2	3	1.1	0.06	9	29 (26)*	15	L	2
(07/18–08/28)	1.0–5.5							L to H	21
1978 (06/10–20)	3.2–4.3	18	3.8	0.25	16	35 (8)*	7	M	3
1980 (06/5–15)	2.5–6.3	45	4.2	0.78				M	4
1981 (10/10–20)	4.0–8.0	36	5.4	1.15	8	55 (11)†		M to H	4
1982 (10/10–20)	9.0–11.5	28	9.8	0.45	6	120 (9)†	20	H	this work
1983 (09/21–25)	2.2–3.3	13	2.8	0.30				L to M	6
1984 (09/17–25)	4.3–5.5	21	4.7	0.22	9	70 (14)†	12	M	5
1985 (04/25–05/02)	2.4–4.2	24	3.4	0.30				M	this work
1987/10–1989/12	1.0–25.0	158	5.7	2.40				L to H	22, 37
Average Etnaean flux			4.0 ± 0.8			35 ± 7‡		M	this work
Average flux from other volcanoes									
Kilauea, Hawaii 1956–83			0.9			3.7			7
Summit degassing Dec. 1983			0.3			3.6			8
Summit degassing Feb. 1984			0.2			1.6			8
East rift eruption Feb. 1984			>10			4.7			8
Merapi, Indonesia 1977–81			0.4			3.3			9
White Island, New Zealand 1985			0.35			1.0			10

All SO₂ fluxes determined by remote correlation spectrometry *n*: number of measurements; s.d.: average standard deviation on the mean flux for a given period. Measurement of the plume (wind) speed by video-anemometry (refs 1–6 and this work) gives a precision of 15–25% on reported SO₂ fluxes. Average CO₂ fluxes were computed from the average SO₂ flux and the average CO₂/SO₂ weight ratio of the plume (given in brackets), corrected for CO₂ air background. Etna fluxes are compared with SO₂ and CO₂ fluxes produced by Kilauea^{7,8}, an oceanic basaltic volcano, and by Merapi (Indonesia)⁹ and White Island (New Zealand)¹⁰, which are representative of steady degassing arc volcanoes.

* CO₂/SO₂ (±5%) measured by gas chromatography after plume sampling in the summit area.

† CO₂/SO₂ measured directly with analysers (Diamant Cosma 6000 infra-red spectrometer for CO₂, Meloy SA 285 flame photometric detector operating at constant pressure for SO₂) during airborne traverses^{4,5}. The standard deviation on CO₂ fluxes includes the contribution of all analytical errors.

‡ See text.

§ The intensity of summit crater activity during the periods of investigation is qualitatively described, on the basis of direct observations, as low (L: no solid ejecta), medium (M: mild strombolian-type activity), and high (H: high strombolian activity). Measurements were contemporaneous with flank lava effusion in 1975, 1976, 1984, 1985, 1987, 1988 and 1989.

and others^{21,22,37} (Table 1) show that this SO₂ flux is variable but generally high, ranging from 1 to 25 Gg (10⁹ g) d⁻¹ as a function of the level of volcanic activity. Our data over a 10-year interval indicate a mean flux of 4 ± 0.8 Gg SO₂ d⁻¹ corresponding to medium volcanic activity. Monitoring by remote correlation spectrometry in 1987–1989 (refs 22, 37) provided a comparable average figure (5.7 ± 2.4 Gg d⁻¹). Such a flux makes Etna an important contributor of volcanic sulphur dioxide to the atmosphere: its average annual release (1.5 ± 0.3 Tg) would approach 10% of the global volcanic output of SO₂ (18.6 Tg yr⁻¹)²³. The emission rate of carbon dioxide was determined by scaling the CO₂/SO₂ ratio of the plume to the corresponding SO₂ flux. Average fluxes ranging from 29 to 120 Gg CO₂ d⁻¹ were obtained, depending on dates and activity level (Table 1). The highest values correspond to high activity (high SO₂ flux), a high plume C/S ratio or both. Average CO₂ fluxes from 35 to 86 Gg d⁻¹ characterize phases of medium activity, the range of which could indicate variable fractionation of CO₂ relative to SO₂ during magma degassing. But because possible SO₂ loss during plume expansion (by adsorption on coarse solid particles and, to a lesser extent, by conversion to sulphate aerosols^{4,6}) could lead to an overestimate of the CO₂/SO₂ ratio, the lower value of 35 ± 7 Gg d⁻¹ is safely considered more representative; it is consistent with a mean SO₂ flux of 4 Gg d⁻¹ and a CO₂/SO₂ weight ratio of ~7, as measured in the hottest (1,090 °C) crater gases. Extrapolating to that value, we estimate 13 ± 3 Tg yr⁻¹ for the mean CO₂ discharge from Etna summit craters in 1975–85. Carbon dioxide emissions from temporary vents during flank eruptions should be added to this budget.

Etna also releases measurable carbon dioxide from its outer flanks, where no visible emissions take place. This emanation was first discovered when a persistent CO₂ excess (20–100 p.p.m.

more than normal atmosphere) was detected^{4,5} at 1.5 m height in the ambient air of the volcano, well outside the plume and far from the craters. This enrichment was verified over a wider region by airborne gas profiling in September 1984 (see Fig. 2). Whereas simultaneous peak enrichment of CO₂ and SO₂ downwind coincides with the visible plume, a CO₂ anomaly (≤10 p.p.m. above the background) persists kilometres away from both sides of the plume, forming a 'dome' of carbon dioxide over the volcano. Given the conditions of the flight, such a signal could not result from simple plume disruption in the lee of the mountain and was computed to arise from a source area ~15 km wide, comprising the volcano flanks above 1,600 m altitude.

Analysis of soil gases in this upper part of Etna²⁴ confirms the occurrence of a wide emanation of carbon dioxide, diluted by air but generally enriched in radon and helium (Table 2). High ground CO₂ contents (from several per cent up to 30% by volume) characterize the cone above ~2,900 m altitude (Piano del Cratere, Fig. 1). This CO₂ has quite uniform δ¹³C of -3.7 ± 0.3‰, similar to that of high-temperature CO₂ from the summit craters^{17,18}, which points to its magmatic derivation. The coexisting helium is of mantle origin, with ³He/⁴He ratio, *R*, reaching 5*R*_a where *R*_a is the atmospheric ratio, or even 8*R*_a, a value typical of mid-ocean-ridge basalts²⁵, when corrected for air dilution. A trend of decreasing CO₂ levels and increasing δ¹³C was observed on the lower slopes (south flank; Table 2). Rather than being evenly distributed, the soil gas anomalies in these areas tend to concentrate along fracture zones (see also refs 26, 27). Recent lava flows usually show no gas anomaly. The CO₂ emanation is thus not due to residual degassing of previously erupted products, but derived from depth. Its δ¹³C reaches values of ~-3.0‰ to -2.0‰ between 2,600 m and 1,800 m altitude; values of up to 0‰, typical of marine carbonates, were

FIG. 1 Sketch map of Etna volcano. The dashed lines enclose the areas where soil gas analysis (226 points) was performed. Numbered points refer to the sampling sites of analyses reported in Table 2. Lava flows from large eruptions between 1984 and 1987 are also drawn (adapted from ref. 35).

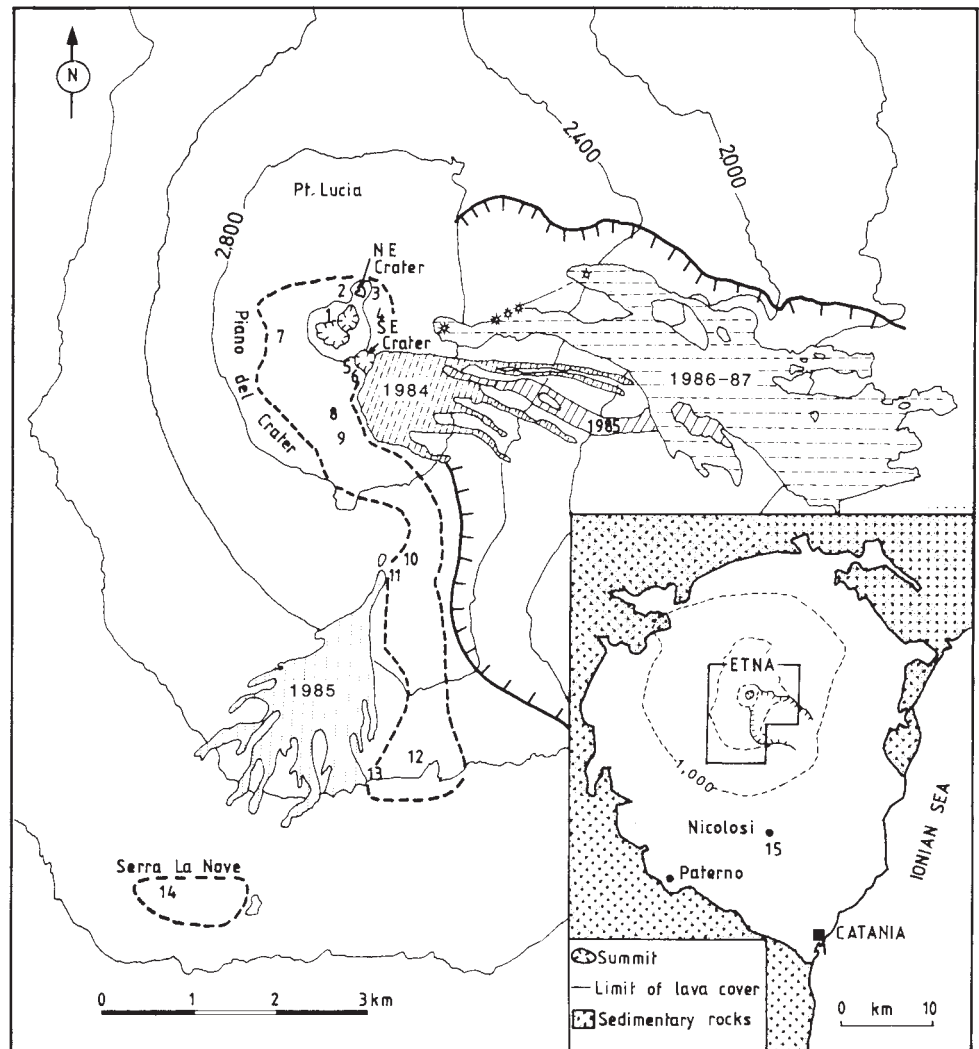
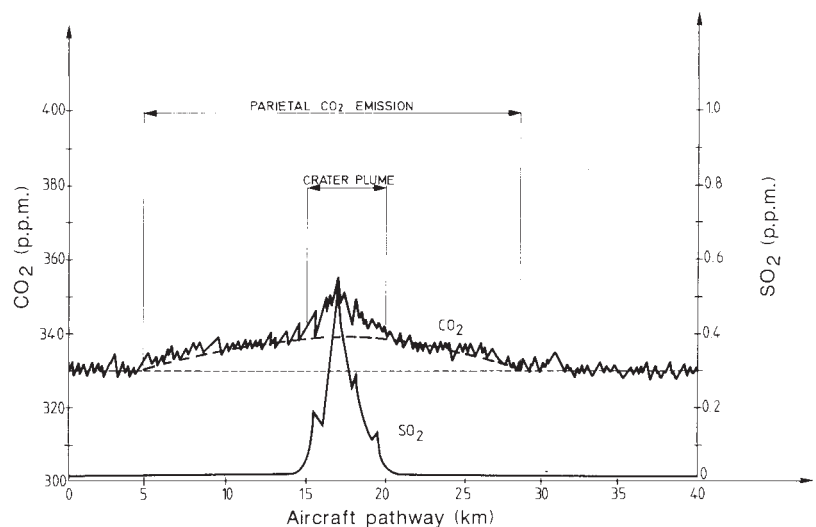


FIG. 2 Airborne horizontal profiling of CO_2 and SO_2 concentrations in the downwind atmosphere of Etna. Data are from a circular flight around Etna made on 22 September 1984, at 3,200 m altitude (plume axis) and 15 km distance from the summit. The wind speed was 8 m s^{-1} and the stable meteorological conditions, corresponding to normal atmospheric diffusion³⁶, constrained the maximum horizontal and vertical plume dispersion from any point source after 15 km ($\sim 1,800 \text{ s}$) transfer to 0.9 km and 0.3 km, respectively. Continuous measurement of CO_2 and SO_2 concentrations was done by infrared spectrometry and flame photometry, respectively, the analysers being periodically calibrated onboard with standard gas. The analytical accuracy is 1 p.p.m. for CO_2 and 10 p.p.b. (parts per 10^9) for SO_2 . Clear peak enrichment of both gases (note the difference in scale) coincides with the visible crater plume. The non-gaussian SO_2 profile is due to degassing from the separate summit craters. The plume width for SO_2 concentration at half maximum value ($\sim 2.5 \text{ km}$) fits well with the total width of the crater source area (1.3 km), plus the diffusional horizontal dispersion at 15 km distance. Of particular interest is the CO_2 anomaly ($\leq 10 \text{ p.p.m.}$) which persists kilometres away from both sides of the crater plume, forming a 'dome' of carbon dioxide over the volcano. Such an extended signal, which cannot result from differential diffusion of the crater CO_2 relative to heavier SO_2 , was computed to originate from Etna's slopes above about 1,600 m altitude, taking account of the horizontal dispersive contribution of sources upwind of the summit crater area (linear distances calculated from the solid angle of the plume). We have verified (see text) that extensive soil degassing of magma-derived CO_2 takes place from this upper part of the cone. A total flank CO_2 output of 55 Gg d^{-1} ($\pm 30\%$) was estimated from these data⁵, considering the mean



CO_2 excess of 5 p.p.m. over a rectangular air section 8 km wide and 1 km high (no vertical concentration gradient being assumed) and multiplying by the wind speed: $8 \times 10^6 (\text{m}^2) \times 8 (\text{m s}^{-1}) \times 1.29 \times 44/29 (\text{kg m}^{-3}) \times 5 \times 10^{-6} = 630 \text{ kg s}^{-1}$. This conservative estimate, probably accurate to within a factor of two, is comparable to the crater plume output ($70 \text{ Gg CO}_2 \text{ d}^{-1}$) measured on the same day.

TABLE 2 Concentration and isotopic composition of CO₂ and He in summit and lateral gas emissions from Etna

Samples	Altitude (m)	CO ₂ (%)	He (p.p.m.)	$\delta^{13}\text{C}$ (‰)	$^3\text{He}/^4\text{He}$ (10 ⁻⁶)	$(^3\text{He}/^4\text{He})_c$ (R/R_a)
Crater gases						
Northeast crater (800 °C)*	3,330	30.8		-4.0		
Northeast crater (1,090 °C)†	3,320	86.7	15.8	-3.6		
Soil gases:						
Piano Bocca Nuova crater (1)‡	3,230	27.9	9.3	-3.7	6.9	8.5
SW flank Northeast crater (2)	3,280	8.0	6.2	-3.5		
E flank Northeast crater (3)	3,275	3.8	5.7	-3.9		
E flank Voragine crater (4)	3,215	5.1	5.6	-3.8		
SW flank Southeast crater (5)	3,100	17.7	7.2	-3.6		
SW flank Southeast crater (6)	3,075	8.1	6.6	-4.1	3.9	8.0
Tabernacolo (7)	3,120	4.2	6.1	-3.7		
N 1971 Osserv. cone (8)	3,000	4.8	5.8	-3.6		
Torre del Filosofo (9)	2,900	1.9	5.7	-3.7		
Piccolo Rifugio 6 (10)	2,530	0.9	5.9	-2.9		
Piccolo Rifugio 11 (11)	2,500	0.7	5.5	-2.6		
Silvestri (12)	2,100	0.35	5.2	-2.3		
Sapienza (13)	2,000	0.33	14.5	-2.5		
Serra La Nave TF 0 (14)	1,800	0.25	6.4	-1.9		
Nicolosi OC (15)	700	0.27		-1.1		
Mofettes						
Paterno 1 (27 °C)	250	96.0	31.1	-0.1		
Paterno 2 (28 °C)	250	98.7	9.5	-0.6	9.2	6.6§
Atmosphere		0.03	5.2	-6.9	1.4	1

All samples were collected in pre-evacuated bottles. Soil gases were collected at 80 cm depth in the volcanic ground, using stainless steel probes filled with a Teflon capillary^{14,24}. Concentration of CO₂ and He, expressed in volume proportions of anhydrous gas, were measured by gas chromatography and mass spectrometry, respectively. Isotopic ratios were measured with double-collecting mass spectrometry, after gas purification under vacuum following routine procedures¹⁸. Carbon isotope results ($\pm 0.1\%$) are given in the usual $\delta^{13}\text{C}$ notation, as $\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{PDB standard}} - 1] \times 10^3$. $^3\text{He}/^4\text{He}$ ratios ($\pm 4\%$) were measured in the Laboratory for Earthquake Chemistry (Tokyo), after separation of He from Ne. They were further corrected for air dilution, using the He/Ne ratio and assuming all neon being atmospheric, to give $(^3\text{He}/^4\text{He})_c$. Because of strong air dilution, the corrected ratios for soil gases have a large uncertainty.

* Post-eruptive venting in August 1979 (ref. 18).

† September 1986 eruption (P.A., unpublished data).

‡ Sampling sites are indicated in Fig. 1.

§ Analysis by B. Marty (MAGIE, France).

measured in CO₂-rich fumaroles near the south base of the volcano (Paterno, Fig. 1). These contain undiluted helium with an R/R_a ratio of 6.6, similar to that for a mofette in the nearby Iblean volcanic complex²⁸, which may thus be typical for basaltic volcanic in Eastern Sicily.

The soil gas data therefore corroborate the airborne data in providing evidence of a diffuse release of igneous CO₂ from the upper flanks of Etna, centred over the axial vents, the amplitude of which decreases as one moves away. Various lines of evidence^{16,29,30} suggest that Etnaean magma batches ascend from a wide reservoir of partial melt at ~20 km depth, through a network of abyssal fissures and dykes. Note that the area affected by soil degassing has an extension comparable to that of the inferred reservoir²⁹. Radial dykes may channel the migration and then the diffuse escape of magmatic volatiles at distances from the main conduits, as revealed by our study. During gas transit and cooling, the most reactive species (such as SO₂) should precipitate into the ground, leaving only carbon dioxide and inert gases to reach the surface, as observed. The increasing $\delta^{13}\text{C}$ of more peripheral CO₂ is consistent with ¹³C-fractionation during deeper (lesser) magma degassing beneath the distal parts of the volcano³¹. Alternatively, it could reflect the growing imprint of CO₂ derived from heating of the carbonate series present in the local basement³², xenoliths of which occur in Etnaean ejecta³³. In any case, our data show that magma-derived carbon dioxide should constitute most of the diffuse flank output. The total diffuse output has been tentatively evaluated using airborne data from the September 1984 profile and taking account of horizontal and vertical dispersion of the corresponding plume during transfer (see caption to Fig. 2). We obtained a conservative estimate of 55 Gg CO₂ d⁻¹ ($\pm 30\%$), which is comparable to the plume output of 70 Gg CO₂ d⁻¹ measured on

the same day⁵. Hence, although more data are needed, we suggest that Etna may release nearly as much CO₂ by diffuse flank emanation as by summit crater degassing. On 22 September 1984 its total discharge probably exceeded 100 Gg CO₂ d⁻¹.

According to these results, Etna is an important emitter of volcanic carbon dioxide. Although it makes up only 0.07% of the annual anthropogenic CO₂ loading (2×10^4 Tg; ref. 34), its CO₂ discharge by sustained crater degassing (13 ± 3 Tg yr⁻¹) is one order of magnitude higher than the annual CO₂ output from Kilauea^{7,8} and actively degassing arc volcanoes^{9,10} (Table 2). Moreover, it represents the flux from the volcanic plume but not from the volcanic pile, as a substantial amount of magmatic CO₂ additionally escapes from the upper cone. Etna is the first volcano where extensive degassing of igneous CO₂ through a volcanic pile—a process anticipated by Sulerzitsky¹⁵ and consistent with a deep release of CO₂ from magmas²⁰—has been documented. Further studies^{11–14} show that such degassing also affects volcanoes in the fumarolic stage. This phenomenon will therefore have to be considered when evaluating the volatile budget of subaerial volcanism and its contribution to the natural background flux of greenhouse gases. As well as being of relevance to volcano monitoring^{13,14}, it is also a possible process for determining radiocarbon ages of vegetation on volcanoes¹⁵ and for CO₂-filling of crater lakes. □

Received 14 February; accepted 5 April 1991.

1. Haultet, R., Zettwoog, P. & Sabroux, J. C. *Nature* **268**, 715–717 (1977).
2. Carbonnelle, J. et al. *Comm. Energ. atom. Rep.* No. 2057, 23 (Paris, 1978).
3. Zettwoog, P. & Haultet, R. *Atmos. Envir.* **12**, 795–796 (1978).
4. Carbonnelle, J. & Zettwoog, P. *Bull. PIRPSEV-CNRS* No. 55 (1982).
5. Carbonnelle, J. et al. *Bull. PIRPSEV-CNRS* No. 108 (1985).
6. Martin, D. et al. *J. geophys. Res.* **91**, 12249–12254 (1986).
7. Gerlach, T. M. & Graeber, E. *J. Nature* **313**, 273–277 (1985).

8. Greenland, L. P., Rose, W. I. & Stokes, J. B. *Geochim. cosmochim. Acta* **49**, 125–129 (1985).
9. Allard, P. et al. *Rep. Volcanol. Investigations in Indonesia*, (CNRS, Paris, 1980).
10. Rose, W. I., Chuan, R. L., Giggensbach, W. F., Kyle, P. R. & Symonds, R. B. *Bull. volcan.* **48**, 181–188 (1986).
11. Badalamenti, B., Gurreri, S., Hauser, S., Parelo, F. & Valenza, M. *Rendiconti Soc. min. petrol. Ital.* **43/4** spec. vol. 893–899 (1989).
12. Allard, P., Baubron, J. C., Luongo, G., Pece, R. & Tedesco, D. *Proc. Int. Conf. Volcanoes*, 212 (Kagoshima, Japan, 1988).
13. Baubron, J. C., Allard, P. & Toutain, J. P. *Nature* **344**, 51–53 (1990).
14. Baubron, J. C., Allard, P., Sabroux, J. C., Tedesco, D. & Toutain, J. P. *J. geol. Soc. Lond.* **148**, 571–576 (1991).
15. Sulerzhitsky, L. D. *Bull. Volcan.* **35**, 85–94 (1971).
16. Chester, D. K., Duncan, A. M., Guest, J. E. & Kilburn, C. R. J. (eds) *Mount Etna, the Anatomy of a Volcano* (Chapman & Hall, London, 1985).
17. Allard, P. *18th IUGG Assembly, IAVCEI Symp. Volcanic Gases* 42 (IUGG, Hamburg, 1983).
18. Allard, P. thesis, Univ. Paris (1986).
19. Faivre-Pierret, R., Martin, D. & Sabroux, J. C. *Bull. volcan.* **43**, 473–486 (1980).
20. Stolper, E. & Holloway, J. R. *Earth planet. Sci. Lett.* **87**, 397–408 (1988).
21. Malinconico, L. L. *Nature* **278**, 43–45 (1979).
22. Caltabiano, T. & Romano, R. *Boll. Grup. naz. Vulcan.* (CNR), 133–145 (1989).
23. Stoiber, R. E., Williams, S. N. & Hubert, B. J. *J. Volcan. geotherm. Res.* **33**, 1–8 (1987).
24. Allard, P. et al. *Terra Cognita* **7**, G17–52, 407 (1987).
25. Craig, H. & Lupton, J. *Earth planet. Sci. Lett.* **31**, 369–385 (1976).
26. Baubron, J. C. *Bur. Rech. Geol. Min. Rep. No. DT-021* (1987).
27. Aubert, M. & Baubron, J. C. *J. Volcan. geotherm. Res.* **35**, 217–225 (1988).
28. Polyak, B. G. et al. *Dokl. Akad. Nauk USSR* **247**, 1220–1225 (1979).
29. Sharp, A. D. L., Davis, P. M. & Gray, F. *Nature* **287**, 587–591 (1980).
30. Armienti, P., Innocenti, F., Petrini, R., Pompilio, M. & Villari, L. *J. Volcan. geotherm. Res.* **39**, 315–327 (1989).
31. Javoy, M., Pineau, F. & Delorme, H. *Chem. Geol.* **57**, 41–62 (1986).
32. Lentini, F. *Memorie Soc. geol. Ital.* **23**, 7–25 (1982).
33. Clochiatti, R., Joron, J. L. & Treuil, M. *Bull. PIRPSEV-CNRS* No. 116 (1986).
34. *US Dept Energy Rep. No. DOE/ER-0239* (Washington, DC, 1985).
35. Armienti, P., Coltelli, M., Innocenti, F., Pompilio, M. & Villari, L. *Boll. Grup. Naz. Vulcan. CNR*, 37–59 (1987).
36. Pasquill, S. (ed.) *Atmospheric Diffusion*, 2nd Edn, 429 (Ellis Horwood, Chichester, 1974).
37. Bruno, N., Budetta, G., Caltabiano, T., Cara, M. & Romano, R. *Boll. Grup. Naz. Vulcan. CNR* (in the press).

ACKNOWLEDGEMENTS. This work was supported by the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique (PIRPSEV, Programme Interdisciplinaire de Recherche pour la Surveillance des Eruptions Volcaniques). We are grateful to Italian colleagues from the International Institute of Volcanology (Catania) and to Etnaean guides for local support, and to H. Wakita (Laboratory for Earthquake Chemistry, Tokyo), D. Tedesco and B. Marty for the isotopic analyses of helium. H. Sigurdsson and T. Trull helped us to improve the manuscript.

The importance of small-scale faulting in regional extension

John Walsh*, Juan Watterson* & Graham Yielding†

* Fault Analysis Group, Department of Earth Sciences, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

† Badley Ashton & Associates Ltd, Winceby House, Winceby, Horncastle, Lincolnshire LN9 6PB, UK

A RECURRING observation in many studies of extensional basins has been that the amount of extension visible on normal faults (for example, on seismic reflection profiles) is significantly less than the amount of extension indicated by crustal thickness and thermal subsidence^{1–5}. One mechanism suggested to account for this discrepancy is small-scale faulting, with offsets too small to be resolved seismically^{6,7}. But earthquake studies^{8–10} indicate that small faults are responsible for only a small fraction of the total seismic moment in an active area. Scholz and Cowie¹¹ have recently attempted to extend this approach to the total strain at the end of a finite deformation interval by combining scaling laws describing the distributions of fault lengths and displacements. Here we present fault displacement data that directly conflict with Scholz and Cowie's conclusions, and imply that up to 40% of the extension may be missed by summing fault offsets on basin profiles. The fault population at the end of a long deformation interval may differ substantially from that responsible for the earthquake population at any one time.

To characterize the fault population in a basin, we have assembled measurements of fault displacements at a wide range of scales: regional seismic data, oilfield seismic data and oilfield well cores. Figure 1a shows cumulative frequency plots of fault displacements on two regional profiles spanning the Viking Graben (northern North Sea); on each profile, the dip-slip displacement of one horizon was measured at all observed faults.

Both of the data curves are roughly linear, indicating a power-law relationship between the fault displacement and the number of readings greater than that displacement. The slopes of the plots are ~ -0.9 and -0.8 . Both data curves become flat at displacements below the effective limit of seismic resolution for these profiles (not all offsets less than ~ 100 m are seen).

Figure 1b shows cumulative frequency curves for fault throws (vertical component of displacement) measured from a detailed depth structure map of the Troll oilfield, located on the regional profiles plotted in Fig. 1a. Throws were measured on parallel sample lines roughly perpendicular to the regional fault strike. Data from a single sample line produce a roughly linear curve like that for the regional profiles in Fig. 1a. As more sample lines are added to the plot, a steep right-hand segment develops, but the central segment maintains a slope between -0.7 and -0.8 , representing the average population slope for the individual sample lines across the oilfield. The overall form of the data curve in Fig. 1b is typical of line samples through faulted areas¹².

Measurements of fault density at the sub-seismic scale are provided by Gabrielsen and Koestler¹³, who examined well cores from the Troll field. In a total of $\sim 1,300$ m of core, the average

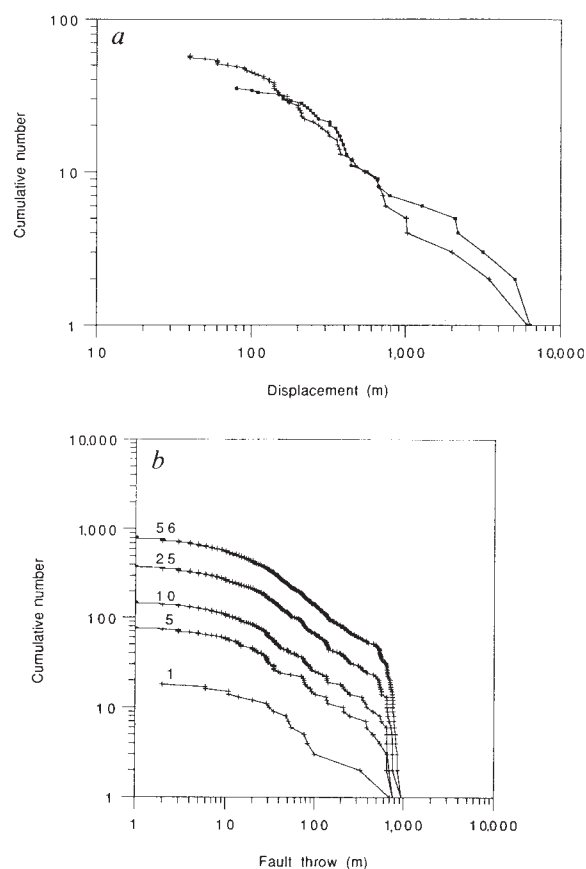


FIG. 1 *a*, Cumulative frequency plots for fault displacements on two regional seismic profiles (profiles I (squares) and II (crosses) of ref. 4). Dip-slip displacements were measured on the uppermost Middle Jurassic on depth-converted interpretations across the north Viking Graben and record the Late Jurassic/earliest Cretaceous episode of faulting^{23,24}. Data from both profiles show a roughly linear distribution on this log-log plot, with slopes of about -0.9 and -0.8 , indicating a power-law relationship between displacement and cumulative number. *b*, Cumulative frequency measurements for fault throws (vertical displacements), measured on sample lines across a depth structure map of a Jurassic horizon in the Troll oilfield (on the regional profiles of *a*). Number of sample lines is indicated for each curve. The plot for a single sample line is roughly linear, like those in (*a*). Adding readings from progressively more sample lines produces a steep right-hand segment, but the central part of the plot maintains the same slope, representing the average population slope on all the sample lines.

fracture density was 0.1 m^{-1} . This figure includes small faults (with displacements down to $\sim 1 \text{ mm}$) and also joints; without fine lithological layering, small faults cannot readily be distinguished from joints.

All of the above datasets comprise displacement measurements on one-dimensional transects through the same extensional basin. The datasets derived from seismic reflection data correspond to sub-horizontal one-dimensional transects (horizon dips being $<10^\circ$, and generally $<5^\circ$). The core data from the Troll field represent vertical one-dimensional transects. The different datasets can be compared with one another by normalizing them in terms of fault density (number of faults intersected per unit distance of sample line). For the core data a correction must be made because the sampling is on vertical transects rather than sub-horizontal transects. Assuming that the seismically resolved faults and the fractures in core have similar dip distributions, the fault density quoted by Gabrielsen and Koestler can be converted to a value equivalent to that which would be observed in a 'horizontal well'. Given the observed range of fault dips, the observed 0.1 m^{-1} in vertical wells would correspond to a horizontally sampled fracture density of $0.09\text{--}0.43 \text{ m}^{-1}$ ($90\text{--}430 \text{ km}^{-1}$).

In Fig. 2a, fault displacement populations from the regional profiles, oilfield maps and core data are all shown on the same plot, with the displacements expressed as heave (horizontal component). Extrapolation of the slope defined by seismic reflection data to displacements at the millimetre scale agrees well with the fault density measurements from the Troll cores. The data are consistent with a single power-law relationship between fault displacements and fault density over six orders of magnitude of fault displacement. The exponent of this power-law relationship (the slope of the data distribution in Fig. 2a) is more negative than -0.7 and may approach -0.9 . A power-law relationship implies that the faulting is scale-invariant (in other words, fractal)¹⁴⁻¹⁷.

Figure 2b shows fault displacement data from another offshore oilfield, taken from well cores and a three-dimensionally migrated (3D) seismic survey. In this case, measurements of cumulative fault density were made in four cored wells. Fault

throws in core were measured, and so cumulative fault density can be estimated at a number of different values of fault throw. The overall slope obtained by treating core and seismic data as one population is about -0.8 .

The observations presented in Fig. 2 imply a power-law relationship of the form $N(d) = ad^{-S}$, where $N(d)$ is the number of fault displacements on the transect having a displacement (or displacement component such as heave or throw) $\geq d$, and a is a constant that depends on the sample size. The exponent S determines the relative contributions of large and small displacements to the total amount of displacement affecting one horizon on the profile.

As bedding dips in the datasets studied are low ($<10^\circ$), we can neglect rotations and consider the heave as being a reasonable measure of the amount of extension on a fault¹⁸. Assuming that the slip direction on the observed faults lies in the plane of the profile, the sum of the heaves represents the fault-related basin extension along the profile. (The total amount of extension may be greater if there is a significant plastic deformation in the rock volume.) Marrett and Allmendinger¹⁹ show that the sum of the displacements on the N largest fault offsets along the one-dimensional transect is given by

$$\sum_{i=1}^N d_i = d_{\max}(1^{-1/S} + 2^{-1/S} + 3^{-1/S} + \dots + N^{-1/S}) \quad (1)$$

and that the total displacement (down to a lower limit of zero displacement) is given by

$$d_{\text{total}} = d_{\max} \left(1^{-1/S} + 2^{-1/S} + \dots + N^{-1/S} + \frac{S(N+1)}{(1-S)(N+1)^{1/S}} \right) \quad (2)$$

which converges rapidly as N is increased. For equation (2) to be finite, $S < 1$. Note that Figs 1 and 2 imply $0.7 < S < 0.9$. By using equations (1) and (2) together, we can consider what fraction of the total displacement is observable above some threshold displacement, to assess the completeness of seismic reflection observations. Figure 3 shows the proportion of total heave seen in a fault displacement population with a maximum

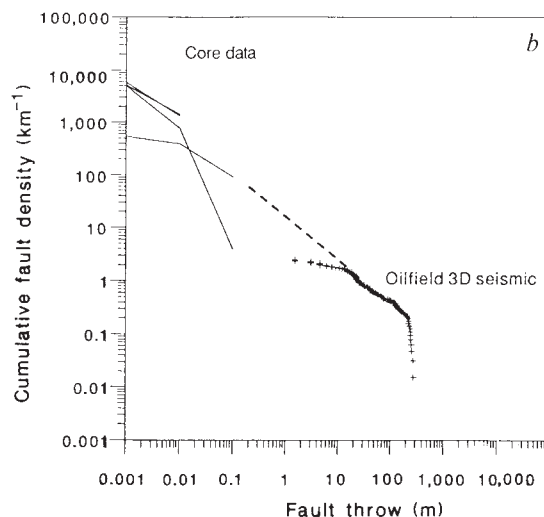
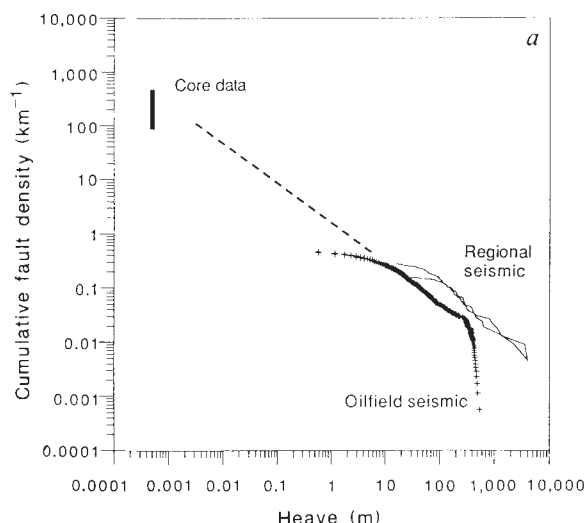
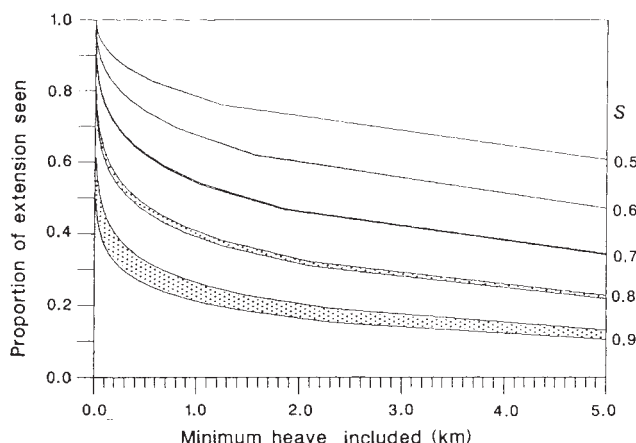


FIG. 2 a, Comparison of fault displacement measurements on regional profiles (solid lines), oilfield seismic mapping (crosses) and oilfield cores (solid bar). Numbers of faults are normalized to cumulative fault density (number of faults per unit length of sample line), and displacements are expressed as heaves (horizontal component). The core data have been corrected to account for the fact that they are from vertical rather than horizontal sample lines; fault dips in the Troll Field have been obtained from depth-converted seismic interpretations (mean $60^\circ \pm 17^\circ$ (2σ)). b, Comparison of fault displacement measurements from three-dimensional

seismic mapping of an oilfield (crosses) and core data from the same oilfield (solid lines). Fault throws were measured in a total of 914 m of core in four different wells (shown as separate lines on the plot). Mapped throws were measured on parallel sample lines in the area around the cored wells. Core measurements have been corrected for the different sampling direction with respect to average fault dip. Despite the broad range of fault density in the wells, overall the measurements are consistent with a single power-law relationship (dashed line) spanning both core and seismic data, with a slope of ~ -0.8 .

FIG. 3 Proportion of fault-related extension seen on a regional profile as a function of the minimum fault heave included in the summation. The maximum heave on the profile is taken as 5 km. Curves are shown for different values of the slope, S , of the fault displacement population. For each value of S , two curves are shown: the upper one assumes that the fault population has a lower cutoff of 1-mm displacement (roughly the sediment grain size), whereas the lower curve assumes that the fault population continues down to zero displacement. Sampling of fault displacements on a regional profile will normally only be complete down to ~ 100 m. For $S=0.8$, only $\sim 60\%$ of the total fault displacement is above this resolution threshold.



heave of 5 km and different values of population slope S , as a function of the minimum heave included in the observations. For whole-basin sections (such as the profiles in Fig. 1a), with the lower limit of resolution 100 m, about 75% of the total displacement is visible for a population slope of $S=0.7$, and only $\sim 60\%$ is imaged for $S=0.8$. Similar calculations for the data in Fig. 2b imply that only $\sim 60\%$ of the fault displacement is seismically imaged. Summing the visible fault heaves on a profile may therefore significantly underestimate the fault-related extension. Similar conclusions are reached by Marrett and Allmendinger²⁰.

To consider the total fault-related strain in the rock volume, fault moments can be summed rather than just displacements^{11,21}. Our observations of fault displacement populations can be converted to moments using the observed scaling relation between length and displacement^{11,19,22}, which has the form $D = cL^n$, where D is the maximum displacement on the fault surface, L is its maximum dimension and c is a constant for a given material property. The exponent n is ~ 1.5 on the basis of a least-squares fit between several hundred values of D and L (refs 19, 22), though Scholz and Cowie¹¹ argue that $n=1.0$. The data presented in Figs 1 and 2 are for fault displacements sampled on one-dimensional transects through the faulted volume; this sample differs from the fault population in three dimensions because large faults are more likely to be intercepted by the sample line than small ones. Marrett and Allmendinger¹⁹ show that, if $N(d) = ad^{-S}$ on the one-dimensional sample line and the faults are self-similar, then in three dimensions

$$N(D) = a_1 D^{-S-2/n} \quad (3)$$

where $N(D)$ is the number of fault surfaces having maximum displacement $\geq D$. If we consider first 'small' faults (smaller than the thickness of the brittle crust), the geometric moment of each fault surface is $M = L^2 D = c^{-2/n} D^{2/n+1}$. Combining this with equation (3) gives the moment sum for all faults

$$\sum_{D=0}^{D_{\max}} M = - \int_0^{D_{\max}} \frac{dN}{dD} M(D) dD = \frac{a_1}{c^{2/n}(S+2/n)(1-S)} D_{\max}^{1-S} \quad (4)$$

The equivalent result for 'large' faults differs only in the value of the constant. Equation (4) converges at a similar rate to equation (1); for example if $D_{\max}$ for an area is 5 km, $S=0.8$ (as Figs 1 and 2), and all faults with $D > 100$ m are included, then only $\sim 55\%$ of the total fault moment would be recognized.

The difference between the results outlined here and those presented by Scholz and Cowie¹¹ arises from the types of data used. Whereas our method is based on objective one-dimensional sampling of fault displacements, which gives a direct measure of extension, the method used by Scholz and Cowie requires an estimate of the fault population in a three-

dimensional volume with fault size expressed as length, where $N(L) = a_2 L^{-C}$. Their proposed value of 1.1 for C is based on two-dimensional sampling of fault trace lengths on maps, and the considerable problems in obtaining the correct exponent by this method are not addressed. A more thorough study¹⁶ of about thirty datasets addresses the differences between one-, two- and three-dimensional population exponents and suggests that $C \approx 3$ for 'small' faults. If this value is used in Scholz and Cowie's method, the results agree with ours.

The observed displacement populations imply that faults with small displacements make a greater contribution to the finite strain than they do to the infinitesimal (seismic) strain. This may seem paradoxical, in that the geological faulting is presumably built up by repeated earthquake cycles, but it is readily understood if the fault system evolves during the period of active faulting. Smaller faults accommodate displacement variations around larger faults, at all scales¹⁴. As faults grow, then even with a constant recurrence interval, many small faults will no longer be required at the positions where they are present and will become inactive. Hence the accumulated geological population of faults will include many more smaller faults than the population that is active at any one time. These possible growth models are explored more fully elsewhere (J.W. and J.W., manuscript in preparation). □

Received 20 December 1990; accepted 12 April 1991.

1. Wood, R. & Barton, P. *Nature* **302**, 134-136 (1983).
2. Ziegler, P. A. *Nature* **304**, 561 (1983).
3. Hellinger, S. J. *et al. Basin Res.* **1**, 191-200 (1988).
4. Marsden, G. *et al. in Tectonic Evolution of the North Sea Rifts* (eds Blundell, D. J. & Gibbs, A. D.) 236-257 (Oxford University Press, 1990).
5. Ziegler, P. A. *Phil. Trans. R. Soc. A* **305**, 113-143 (1982).
6. Angelier, J. & Colleta, B. *Nature* **301**, 49-51 (1983).
7. Kautz, S. A. & Solater, J. G. *Tectonics* **7**, 823-832 (1988).
8. Gutenberg, B. & Richter, C. F. *Seismicity of the Earth and Associated Phenomena* (Princeton University Press, New Jersey, 1954).
9. Wyss, M. *Geophys. J. R. astr. Soc.* **31**, 341-359 (1973).
10. Jackson, J. & McKenzie, D. *Geophys. J.* **93**, 45-73 (1988).
11. Scholz, C. H. & Cowie, P. A. *Nature* **346**, 837-839 (1990).
12. Childs, C. *et al. in North Sea Oil and Gas Reservoirs II*, 309-318 (Graham & Trotman, London, 1990).
13. Gabrielsen, R. H. & Koestler, A. G. *Norsk geol. Tidsskr.* **67**, 371-381 (1987).
14. King, G. C. P. *Pure appl. Geoph.* **121**, 761-815 (1983).
15. Turcotte, D. *Tectonophysics* **132**, 261-269 (1986).
16. Heffer, K. & Bevan, T. in *Proc. 2nd Eur. Petrol. Conf.* 367-376, Soc. Petrol. Eng. reprint No. 20981 (1990).
17. Sassi, W. *et al. in Structural and Tectonic Modelling and its Application to Petroleum Geology* (ed. Norsk Petroleumsforening) (Graham & Trotman, London, in the press).
18. Solater, J. G. & Celerier, B. *Basin Res.* **1**, 217-221 (1989).
19. Marrett, R. & Allmendinger, R. W. *J. struct. Geol.* (in the press).
20. Marrett, R. & Allmendinger, R. W. *Geology* (in the press).
21. Kostrov, V. V. *Izv. Earth Physics* **1**, 13-21 (1974).
22. Walsh, J. & Watterson, J. *J. struct. Geol.* **10**, 239-247 (1988).
23. Badley, M. E. *et al. J. geol. Soc.* **145**, 455-472 (1988).
24. Yielding, G. *et al. Spec. Publ. geol. Soc. Lond.* (in the press).

ACKNOWLEDGEMENTS. We thank the Troll Field partners, especially Statoil, that have provided seismic and core data for fault-population analysis, and C. Lavers, P. Mason and M. Seymour for assistance with that analysis. Discussions with B. Freeman, A. Roberts and M. Casey are gratefully acknowledged. We thank T. Bevan, K. Heffer and R. Marrett for preprints of their work, and P. England for a constructive review of an earlier version of this paper.

Close tetrapod relationships of the coelacanth *Latimeria* indicated by haemoglobin sequences

Thomas Gorr, Traute Kleinschmidt* & Hans Fricke†

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, Germany

† Max-Planck-Institut für Verhaltensphysiologie, D-8131 Seewiesen, Germany

* To whom correspondence should be addressed

THE origin of tetrapods has been debated for many years. In traditional systematics, the extinct lobe-finned bony fish (Rhipidistia) are regarded as the closest relatives of tetrapods¹. Among living fish, the coelacanth *Latimeria chalumnae* (Actinistia)²⁻⁴, which is the only recent representative of the Crossopterygii (Actinistia and Rhipidistia), the lungfish (Dipnoi)⁵⁻⁸ and ray-finned fish (Actinopterygii)^{9,10}, have each been considered as sister-groups of the tetrapods. We have now determined the sequence of the α - and β -globin chains of coelacanth haemoglobin and compared them with all known haemoglobins of bony and cartilaginous fish as well as those of tadpoles and adult amphibians. Haemoglobins of bony fish match more closely those of larval than adult amphibians. The β chains of *Latimeria* match those of tadpoles more closely (54%) than do those of any other fish, whereas the α chains of *Latimeria* (45.4%), and especially of teleosts (49.2%), are closer to those of larval amphibians than are those of lungfish (39.8%). If only synapomorphic sequence matches (those at derived positions shared by one bony fish and tadpoles but not by

any other bony fish) are considered, both *Latimeria* globin chains have distinctly more identities with phase of tadpoles than do those of any bony fish. Thus the primary structure of *Latimeria* haemoglobin indicates that the coelacanth is the closest living relative of tetrapods.

Taxonomic relationships can be reconstructed by comparison of aligned homologous proteins such as haemoglobin, for which many amino-acid sequences are available¹¹. *Latimeria* haemoglobin chains were sequenced by Edman degradation (T.G., T.K., J. G. Sgoñros and L. Kasang, manuscript in preparation), and the primary structures of α and β chains are shown in Fig. 1 with the amino-acid sequences of haemoglobins from a selection of bony fish, including a lungfish.

To select chains giving the most matches to *Latimeria* α and β chains, each chain was compared with a pool of sequences stored in the database MIPSX using the program FASTA¹². Among the 522 haemoglobins were six chains of cartilaginous fish (3 α , 3 β), two of lungfish (1 α , 1 β), 12 of modern bony fish (teleosts, 6 α , 6 β), 11 of adult tailless amphibians (anurans, 5 α , 6 β), 10 of larval anurans (5 α , 5 β) and nine of tailed amphibians (urodelans, 7 α , 2 β). Of all 522 haemoglobins, the β chains of coelacanth and two β -globin chains of the frog *Rana catesbeiana* tadpoles had the highest match with 58% identity (optimized score 499), separated from other comparisons by over 5 standard deviations above the mean score. Because neither fish nor adult amphibian sequences showed similar matches to *Latimeria*, we did a Needleman-Wunsch comparison¹³ in pairs including tadpoles together with all the above species. Table 1 shows that the β -globin chains of teleosts, *Latimeria* and *Lepidosiren* are more similar to those of larval than adult amphibians, as are the α -globin chains of at least teleosts and *Latimeria*.

The close match to tadpole globins may be explained by the evolutionary relationships of the genes involved. DNA sequenc-

TABLE 1 Needleman-Wunsch comparison of globin chains from adult and larval amphibians with those of fishes

α -CHAIN COMPARISON: IDENTICAL MATCHES (%)												
	Anura adult					Anura larval					Urodela	
	Xen. laev.	Xen. laev.	Xen. bor.	Xen. bor.	Xen. trop.	Xen. laev.	Xen. laev.	Xen. laev.	Xen. trop.	Rana cat.	Pleu. walt.	Trit. crist.
Amphibians:	α_M	α_m	$\alpha-l$	$\alpha-ll$	α	$\alpha-T3$	$\alpha-T4$	$\alpha-T5$	α	$\alpha-III$	α_M	α_m
Fishes												
Heterodontus- α	42.6	42.6	43.6	41.4	42.9	42.9	42.1	40.0	41.8	38.6	40.7	40.4
Lepidosiren- α	40.4	39.7	40.7	40.0	42.1	40.7	40.7	40.7	39.0	37.9	38.6	38.3
<i>Latimeria</i> - α	42.6	42.6	43.3	43.3	47.5	45.4	44.7	46.1	42.6	48.2	46.8	43.3
Salmo I- α	48.2	46.8	47.5	49.7	48.9	56.7	54.6	57.5	51.8	55.3	48.9	50.4
Cyprinus- α	41.8	41.1	42.6	39.7	40.4	48.9	46.8	46.8	45.4	47.5	44.7	46.8
Electrophorus- α	44.7	45.4	45.4	44.0	44.7	48.9	46.8	45.4	45.4	51.8	46.1	46.8
Catostomus- α	39.0	39.0	39.7	37.6	39.0	46.1	44.0	47.9	43.3	46.8	41.8	44.7
Thunnus- α	45.8	43.7	46.8	44.0	46.1	53.9	52.5	51.8	49.3	44.7	44.7	45.8
β -CHAIN COMPARISON: IDENTICAL MATCHES (%)												
	Anura adult					Anura larval					Urodela	
	Xen. laev.	Xen. bor.	Xen. trop.	Rana esc.	Rana cat.	Xen. laev.	Xen. laev.	Xen. trop.	Rana cat.	Rana cat.	Trit. crist.	Trit. crist.
Amphibians:	β_M	$\beta-l$	β	β	β_M	$\beta-1$	$\beta-2$	β	$\beta-III$	$\beta-IV$	β_M	β_m
Fishes												
Heterodontus- β	32.1	32.1	37.6	37.8	33.3	44.0	44.0	45.4	45.4	45.4	32.9	31.4
Lepidosiren- β	44.1	44.1	45.9	45.0	41.4	47.3	45.2	45.9	48.6	49.3	44.1	40.7
<i>Latimeria</i> - β	42.1	41.4	43.2	40.7	40.7	52.1	50.0	52.1	58.2	57.5	49.0	42.1
Salmo I- β	38.6	39.3	39.0	45.7	42.9	53.4	49.3	52.7	52.1	50.7	40.7	38.6
Salmo IV- β	40.7	40.7	46.6	41.4	39.3	48.6	44.5	46.6	52.1	49.3	46.2	44.1
Cyprinus- β	43.5	42.8	44.5	45.0	42.9	47.3	43.8	45.9	49.3	49.3	42.8	43.5
Electrophorus- β	44.8	44.8	45.9	48.6	42.1	46.6	41.8	43.8	49.3	48.0	45.5	42.8
Thunnus- β	40.0	39.3	43.8	44.3	42.9	46.6	44.5	43.8	48.0	46.6	40.7	41.4

Investigated species: Anurans adult and anurans larval: Xen. laev., *Xenopus laevis*; Xen. bor., *Xenopus borealis*; Xen. trop., *Xenopus tropicalis*; Rana esc., *Rana esculenta*; Rana cat., *Rana catesbeiana*. Urodela: Pleu. walt., *Pleurodeles waltlii*; Trit. crist., *Triturus cristatus*. Cartilaginous fishes: *Heterodontus portusjacksoni*. Lungfishes: *Lepidosiren paradoxus*. Crossopterygians: *Latimeria chalumnae*. Actinopterygians represented by teleosts: *Salmo irideus*; *Cyprinus carpio*; *Electrophorus electricus*; *Catostomus clarki*; *Thunnus thynnus*.

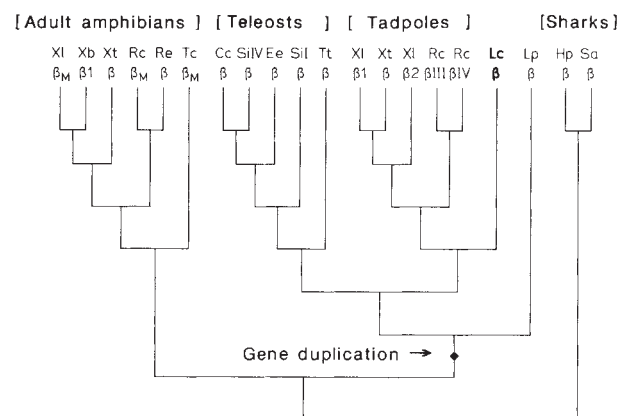


FIG. 2 Clustal β tree. Species included: XI β_{M} , *Xenopus laevis* (frog, β_{maj}); Xb β_1 , *Xenopus borealis* (frog, β_1); Xt, *Xenopus tropicalis* (frog); Rc β_M , *Rana catesbeiana* (frog, β_{maj}); Re, *Rana esculenta* (frog); Tc β_M , *Triturus cristatus* (newt, β_{maj}); Cc, *Cyprinus carpio* (β -A); SiIV, *Salmo irideus* IV; Ee, *Electrophorus electricus*; SiI, *Salmo irideus* I; Tt, *Thunnus thynnus*; XI β_1 , *Xenopus laevis* (tadpole, β_{T1}); Xt, *Xenopus tropicalis* (tadpole); XI β_2 , *Xenopus laevis* (tadpole, β_{T2}); Rc β_{III} , *Rana catesbeiana* (tadpole, β_{III}); Rc β_{IV} , *Rana catesbeiana* (tadpole, β_{IV}); Lc, *Latimeria chalumnae*; Lp, *Lepidosiren paradoxus*; Hp, *Heterodontus portusjacksoni*; Sa, *Squalus acanthias*.

teleost α chains next to the tadpoles leaving either *Latimeria* or *Lepidosiren* as out-groups. Only one consistent order of species was obtained for the β chains using all three methods: (((*Latimeria*-tadpoles) teleosts) *Lepidosiren*) frogs. This order did not vary with the number of species included (up to 20 with 26 chains).

For an equivalent result for both chains, we used the cladistic concept²¹ which evaluates sister-group relationships using only those characters exclusively shared between different species and derived from their most recent ancestor (synapomorphies²¹). Primitive characters that are retained from a very ancient progenitor (symplesiomorphies^{21,22}) should be excluded. Adult and

larval anuran chains had a high number of symplesiomorphous identities which persist from before the adult-larval duplication in both gene lineages rather than being convergently identical. The degree of conservation of these residues in fish globin chains reflects their rates of mutation and the overall high number of matches between teleost α chains and those of amphibians is partially based on a high conservation of these plesiomorphous residues (Table 2, column 1). This indicates that the younger teleost α chains have a much lower mutation rate than do the ancient α chains of *Latimeria* and *Lepidosiren* and that the teleost-tadpole cluster, revealed by the most parsimonious method of Goodman *et al.*^{10,15}, as well as by the CLUSTAL

TABLE 2 Bony fish/tadpoles weighted homologies

		α -chain comparison					
Post-duplication state: orthologous positions (OPs)							
Bony-fish species	Preduplication state: symplesiomorphous positions	OPs/species	Synapomorphous positions (SAPs) per:				
			Teleosts or sarcopteryg.	Species	SAPs per species		
Latimeria- α	27	18	→	10	→	8	= α -12S/13I/21T/32L/55L/71N/76I/131E/
Lepidosiren- α	18	12	→	7	→	5	= α -2F/63V/78S/82T/104A/
Cyprinus- α	22	12	→	17	→	4	= α -8A/9A/19P/115G/
Electrophorus- α	29	14	→	14	→	3	= α -19S/68E/82S/
Salmo I- α	33	18	→	13	→	5	= α -19G/60G/96G/111I/134A/
Thunnus- α	25	13	→	16	→	5	= α -1T/71T/115K/120D/121A/
β -chain comparison							
Latimeria- β	31	30	→	16	→	15	= β -9A/16Q/29T/32F/51S/55A/76H/86H/91K/95E/104K/111I/131E/133L/142S/
Lepidosiren- β	31	18	→	9	→	8	= β -13S/16S/19D/22H/35F/71I/84H/85L/
Cyprinus- β	21	11	→	7	→	1	= β -123S/
Electrophorus- β	22	14	→	11	→	5	= β -79N/90T/91I/108E/128H/
Salmo I- β	25	21	→	12	→	7	= β -20I/86K/93A/94N/97F/120A/146F/
Thunnus- β	19	14	→	8	→	5	= β -12A/21E/58A/86S/121A/

This reveals the numbers of matches when these are separated between shared ancestral (symplesiomorphous) and exclusively shared derived matches (synapomorphous). Only the latter ones can evaluate relationships in a cladistic sense. Column 1 shows an unexpectedly high number of symplesiomorphous positions found in the α chains of the more recent teleosts with a fossil record of 220 Myr³⁰, a lower one for *Latimeria* and lower still for the lungfish *Lepidosiren*, compared with their β chains. In agreement with Goodman *et al.*¹⁷ this points to the high mutation rate of early α chains, particularly that of *Lepidosiren*. Note the correlation between most symplesiomorphies found in *Salmo* α chains and the highest number of matches to amphibian α chains in Table 1. That the ancient haemoglobins of *Latimeria* and *Lepidosiren* (fossil records of coelacanths aged 380 Myr and of dipnoans aged 400 Myr^{5,30}) better conserved these ancestral positions in their β chains, indicates a homogeneous rate of evolution for these subunits and confirms that *Latimeria* corresponds more closely to tadpoles than do the teleosts (Table 1). Column 2 shows, for both chains, the number of orthologous positions between all tadpoles and each bony-fish species, excluding positions shared by tadpole and human haemoglobin (α , β) as a paralogous sequence (; in Fig. 1). Of the orthologous positions the number of synapomorphous positions (SAPs) were counted. Depending on the different numbers of sequences for teleosts (4), *Latimeria* (1) and *Lepidosiren* (1), there is not only a four-times higher competition between the teleosts and *Latimeria* or *Lepidosiren*, but also a three-times higher competition between the teleosts themselves if only SAPs per fish species are considered. To examine this increased intra-group competition another counting mode (column 3) treats the teleosts or sarcopterygians (*Latimeria* and *Lepidosiren*) as unities for the determination of the SAPs specific for each group to tadpoles. Column 4 shows the number of SAPs for each bony-fish species with all tadpoles. As *Latimeria* represents the most SAPs in both chains with tadpoles among all considered fish species, these unique matches support the results for the β chains of pairwise comparisons (Table 1) and of both phylogeny programs (Fig. 2). Column 5 lists all SAPs for each included bony fish species (see Fig. 1).

and PROTPARS programs, seems to be a paraphyletic grouping.

Real sister-groups can be established using synapomorphies in orthologous sequences^{16,23}. For α chains there were eight synapomorphies for *Latimeria* compared with five putative ones for *Lepidosiren*, and three to five for the teleosts (Table 2, fourth column). The β chains had 15 synapomorphies for *Latimeria* versus eight putative ones for *Lepidosiren*, and up to seven for the teleosts.

As the risk of paraphyletic grouping is evident for the α chains, and the probability of misleading results increases with high and unequal mutation rates in different lineages^{24,25}, we present only a β -chain tree (Fig. 2) calculated by CLUSTAL for 14 species (20 chains). It reflects the sister-group relationship

of *Latimeria* and tadpoles joined by the out-group of teleosts and then *Lepidosiren*. The pairwise comparison and the tree reconstructions show that globin chains of fish are orthologous to those of tadpoles and not to those of frogs. Adult bony fish express haemoglobin chains of larval type, which is an example of the retention of larval characters into adult life (paedomorphosis).

In agreement with traditional systematics¹ and recent anatomical investigations^{2,3,26}, as well as studies comparing 28S ribosomal DNA sequences²⁷ and myelin proteins²⁸, these data indicate that *Latimeria* is the closest extant relative of tadpoles (representing tetrapods) with respect to both α - and β -globin chains. Furthermore, many other reported comparisons group the coelacanth closer to tetrapods than any lungfish. □

Received 24 October 1990; accepted 15 April 1991.

1. Romer, A. S. *Vergleichende Anatomie der Wirbeltiere* 4th edn, 37–88 (Parey, Hamburg, Berlin, 1976).
2. Schultze, H.-P. *J. Morph. Suppl.* **1**, 39–74 (1986).
3. Long, J. A. *J. Vertebr. Paleont.* **9**, 1–17 (1989).
4. Nelson, J. S. *Fishes of the World* 2nd edn (Wiley, New York, 1984).
5. Benton, M. J. *J. molec. Evol.* **30**, 409–424 (1990).
6. Panchen, A. L. & Smithson, T. R. *Biol. Rev.* **62**, 341–438 (1987).
7. Rosen, D. E., Forey, P. L., Gardiner, B. G. & Patterson, C. *Bull. Am. Mus. nat. Hist.* **167**, 159–276 (1981).
8. Forey, P. L. *J. Morph. Suppl.* **1**, 75–91 (1986).
9. Wahlert, G. von *Latimeria und die Geschichte der Wirbeltiere* (Fischer, Stuttgart, 1968).
10. Goodman, M., Miyamoto, M. M. & Czelusniak, J. in *Molecules and Morphology in Evolution: Conflict or Compromise?* (ed. Patterson, C.) 141–176 (Cambridge University Press, Cambridge, 1987).
11. Kleinschmidt, T. & Sgouros, J. G. *Biol. Chem. Hoppe-Seyler* **368**, 579–615 (1987).
12. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
13. Needleman, S. B. & Wunsch, C. D. *J. molec. Biol.* **48**, 443–453 (1970).
14. Hosbach, H. A., Wyler, T. & Weber, R. *Cell* **32**, 45–53 (1983).
15. Goodman, M., Romero-Herrera, A. E., Dene, H., Czelusniak, J. & Tashian, R. E. in *Macromolecular Sequences in Systematic and Evolutionary Biology* (ed. Goodman, M.) 115–191 (Plenum, New York, 1982).
16. Fitch, W. *Syst. Zool.* **19**, 99–113 (1970).

17. Goodman, M., Moore, G. W. & Matsuda, G. *Nature* **253**, 603–608 (1975).
18. Higgins, D. G. & Sharp, P. M. *Gene* **73**, 237–244 (1988).
19. Felsenstein, J. *Evolution* **39**, 783–791 (1985).
20. Felsenstein, J. *Phylog. 3.2 Manual* (University of California Herbarium, Berkeley, California, 1989).
21. Hennig, W. *Phylogenetic Systematics* (University of Illinois Press, Urbana, Chicago, London, 1966).
22. Ax, P. *Das Phylogenetische System* (Fischer, Stuttgart, New York, 1984).
23. Dobzhansky, T., Ayala, F. J., Stebbins, G. L. & Valentine, J. W. *Evolution* 262–313 (Freeman, San Francisco, 1977).
24. Felsenstein, J. *Syst. Zool.* **27**, 401–410 (1978).
25. Felsenstein, J. A. *Rev. Ecol. Syst.* **14**, 313–333 (1983).
26. Fritzsche, B. *Nature* **327**, 153–154 (1987).
27. Hillis, D. M. & Dixon, M. T. in *The Hierarchy of Life-Molecules and Morphology in Phylogenetic Analysis* 355–367 (Excerpta Medica, Elsevier, Amsterdam, 1989).
28. Waehndt, T. V. & Malotka, J. *J. Neurochem.* **52**, 1941–1943 (1989).
29. Watt, K. W. K. & Riggs, A. *J. biol. Chem.* **250**, 5934–5944 (1975).
30. Müller, A. H. *Lehrbuch der Paläozoologie* 2nd edn Vol. 3/1, 194–443 (Fischer, Jena, 1985).

ACKNOWLEDGEMENTS. We thank L. Kasang for *Latimeria* erythrocytes, M. de San (FED-Project) for assistance, R. Rossi for laboratory facilities, B. Schrank for gas- and liquid-phase sequencing, J. G. Sgouros and A. Elzanowski for helpful discussions and C. Oubridge and T. H. Jensen for constructive criticism of the manuscript.

Relation of an array of early-differentiating cones to the photoreceptor mosaic in the primate retina

Kenneth C. Winkler & Pasko Rakic

Section of Neurobiology, Yale University School of Medicine, PO Box 3333, New Haven, Connecticut 06510, USA

THE retina of diurnal primates, including humans, contains a reiterative mosaic of red-, green- and blue-sensitive cones whose visual pigments are maximally sensitive to long, middle or short wavelengths, respectively¹. Although the distribution of the cone subtypes in the adult rhesus monkey has been quantified using opsin-specific antisera², the mechanism for the phenotypic specification of the cone subtypes and the establishment of their ratios in the retinal mosaic remain unknown. Here we present immunocytochemical evidence that a subset of cones (about 10%) express their cell-specific opsin two to three weeks before the surrounding cones. Remarkably, these precocious cones are evenly stationed throughout undifferentiated regions of the retinal surface from several weeks after their last mitotic division³, and at least one month before the formation of their synapses with bipolar and horizontal cells⁴. Use of confocal laser microscopy reveals that the inner segments of immunolabelled and surrounding unlabelled cones are transiently in apposition with one another, enabling surface mediated interactions to occur during this period. We suggest that the early maturing cones induce neighbouring undifferentiated cones to express an appropriate opsin phenotype, and therefore constitute a 'protomap' for the emergence of the species-specific retinal mosaic.

We examined the emergence of the cone-specific photopigments in the retinal cell mosaic using antisera raised against synthetic peptides after cloning and sequencing complementary DNAs that encode human photopigment polypeptides^{5,6}. The presence of the cone pigments was detected using antibodies specific for either the red/green or blue opsins^{2,7}. The 96% sequence homology between the red and green opsins prevents them from being distinguished at present. Retinal whole-mounts were prepared from 12 rhesus monkeys (*Macaca mulatta*)—10 fetuses, killed at embryonic (E) day 65, 75, 80, 90, 100, 110, 120, 130, 140 or 165 (the day of birth) and 2 normal adults. The retina from one eye was immunolabelled with antibody 4942A which recognizes both the red- and green-sensitive opsins, and the other eye was immunolabelled with antibody 108B specific for the blue-sensitive opsin^{2,7}. In adults, 85–90% of all cones are immunoreactive for the 4942A antibody and 10–15% for the 108B antibody (Fig. 1a). The mosaic of rods and cones exhibits a regular arrangement across the entire surface of the adult retinae^{2,8}.

To examine the emergence of this rod-cone mosaic, we first analysed unstained fetal retinal whole-mounts using differential interference contrast (DIC) optics. At all ages examined, cones could be distinguished from rods by their larger size (Fig. 1b). But from E65 to E120, photoreceptors were arranged in an adult-like mosaic only in the central retina. At these ages, cone inner segments in the periphery were not yet separated from each other by rod inner segments (Fig. 1, compare b and c). Indeed, confocal microscopy of an E90 retinal whole-mount immunoreacted with the monoclonal antibody CSA-1 specific for cone membranes, suggested that the surfaces of cone inner segments are transiently in apposition with one another (Fig. 2a). The paucity of rod inner segments in immature regions of the retina was expected because the genesis and differentiation of rods in a given sector lags behind that of cones by as much as one month^{3,9}.

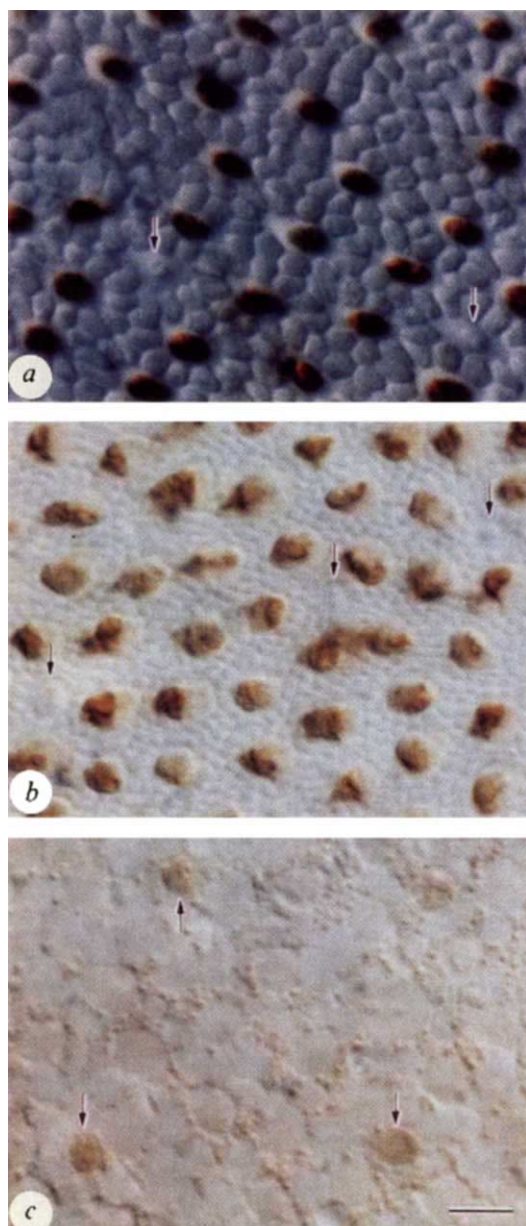


FIG. 1 *a*, Red- and green-sensitive cones in the adult monkey retina labelled by the 4942A antibody. Retinae were marked for orientation, dissected from other tissues, and mounted flat on a slide. The whole-mount was then processed for immunocytochemical visualization of the 4942A polyclonal antibody as described². The larger photoreceptors are cones and the smaller objects surrounding the cones are rods. About 90% of cones across the retina were 4942A-positive, as indicated by the brown reaction product. The unlabelled 10% of cones (arrows) can be labelled by 108B, a blue opsin-specific polyclonal antibody². Nonspecific staining was assessed at each age in preparations that were processed in parallel, with omission of the primary antisera. No labelling of photoreceptors was evident in these controls. *b*, The topography of 4942A-labelled cones in the central retina of a 90-day-old monkey fetus. The distribution and morphology of red/green-sensitive cones are similar to that of the adult. In this region, as in mature retinae, about 10% of all cones are unlabelled by the 4942A antibody (some examples indicated by arrows). As is evident in *a*, the smaller objects surrounding the cones are rods. *c*, The periphery of the same retina as in *b*, immunoreacted with 4942A, shows only three labelled cones (arrows), whereas the numerous unlabelled cones are present in the surrounding mosaic. Eventually, 90% of the cones (large profiles) will become 4942A-positive. Only a few rods are seen in this photograph. They will eventually grow into the mosaic and separate the cones from each other as illustrated in *a* and *b*. Scale bar, 10 μ m.

Examination of E65 and E75 retinae immunoreacted with opsin-specific antisera failed to reveal labelled cones. The first red/green-sensitive (4942A-positive) cones were found in the fovea and perifoveal region of the E80 specimen. Blue-sensitive (108B-positive) cones, by contrast, were first encountered about three weeks later, between E100 and E110. At E90, the distribution of red/green-sensitive cones in a 50-mm² area centred at

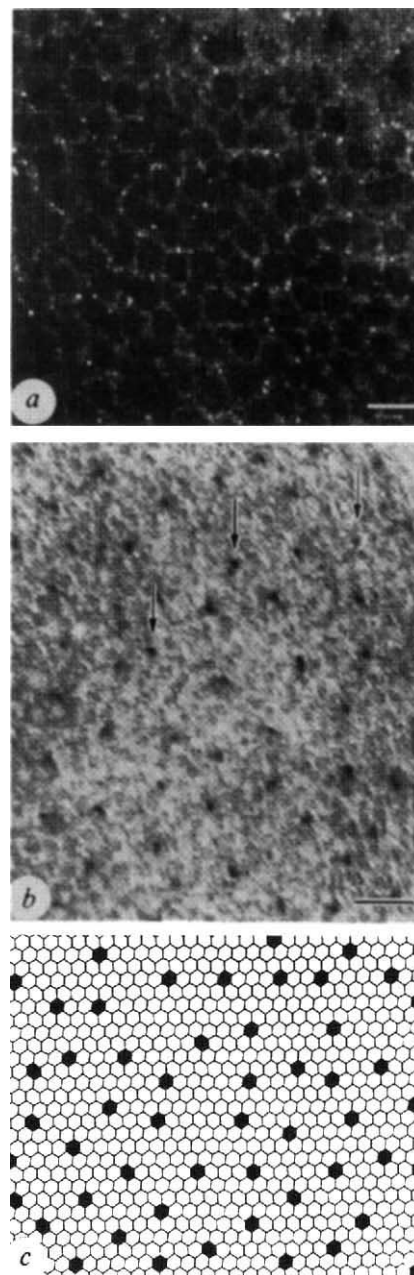


FIG. 2 *a*, Confocal laser microscopic section through cone inner segment membranes immunoreactive to the CSA-1 antibody in a 90-day-old fetal monkey retina. CSA-1 is a monoclonal antibody specific to cone inner and outer segment membranes²⁰. Although cones are separated from each other by the rods in adult retinae (see Fig. 1*a, b*), cone membranes are in apposition with one another in the retinal periphery between E80 and E110. Scale bar, 10 μ m. *b*, A video-enhanced, DIC optics photograph of red/green-sensitive (4942A-positive) cones in the periphery of an E80 retina. Cones that precociously express the red/green opsin are positioned in a regular array that is distinct from their arrangement in the adult photoreceptor mosaic. Scale bar, 20 μ m. *c*, Schematic diagram of the retinal surface at the periphery of the E90 specimen illustrating the distribution of 4942A-labelled cones (black) among the densely packed mosaic of unlabelled cones.

the fovea resembled that in the adult (Figs 1b and 3). In this central region, about 90% of all cones were immunoreactive to the red/green opsin-specific antibody, as in the adult. Outside this perifoveal region, the frequency of cones that expressed the red/green opsin was low (about 10% of all cones in a given area) and the ratio of labelled to unlabelled cones appeared to be the inverse of the pattern in central retina (Figs 1c and 2b, c). These 4942A-immunoreactive cones were observed in the peripheral rim in all fetuses from E80 to E110, although they generally were labelled less intensely than cones in the central retina. They were spaced in a regular periodic pattern (Fig. 2b, c): each immunolabelled cone in the peripheral retina was surrounded by a set of immature cone segments which did not react with either the red/green opsin-specific or the blue opsin-specific antisera.

The extent of the regions occupied by different ratios of 4942A-labelled and unlabelled cones changed with age (Fig. 3). Thus, between E80 and E110, the central region of the retina containing 4942A-labelled and unlabelled cones in the adult ratio of 10 to 1 gradually spread towards the periphery. Concomitantly, the area containing cones that precociously express this opsin at a ratio of 1 to 10, became progressively restricted to the retinal margins. A region containing an intermediate ratio of labelled to unlabelled cones averaging 1 to 1 was found between the mature area of labelled cone density and the area occupied by the mosaic of precocious cones (Fig. 3). This region may represent a transition zone from the immature array to the adult-like mosaic of cone subtypes. As unlabelled cones begin to express their opsin and rod outer segments enter the mosaic, the array of precocious cones was replaced by the adult pattern. Therefore, the centre-to-periphery gradient of opsin-specific expression parallels the topographic development determined from ^3H -labelled thymidine autoradiographic analyses of the time of cell origin in the mammalian retina^{3,9-11}. The adult mosaic occupied the entire retina surface by E130. Significantly, precocious cones are immunoreactive to the red and green opsin-specific antibody at least one month before the formation of synaptic contacts between photoreceptors and either horizontal or bipolar cells⁴. Therefore, the specification of cone subtypes is not induced by synaptic communication with other retinal cell types.

The present results indicate that preceding the centre-to-periphery maturational wave of opsin expression, there is a zone of predominantly unlabelled cones that contains a small subset of earlier-maturing red/green-sensitive cones. What is the role of cones that precociously express the red/green opsin? Do early-differentiating cones constitute a 'protomap' for the setting up of repeating areal units of about ten red/green- and one blue-sensitive cone?

We can envision at least two scenarios in which precocious cells specify the composition of the primate cone mosaic. One hypothesis is that the precocious cones are prospective blue-sensitive cones which initially express the red/green opsin, and later switch to the blue opsin. This is suggested by the similarity in the spacing of precocious red- and green-sensitive cones in fetal retinae and blue-sensitive cones in the adult mosaic. An alternative hypothesis, which we favour, is that early-differentiated cones emit signals either by making contact with adjacent undifferentiated cones or through diffusible agents as proposed for the development of ommatidia in the compound eye of *Drosophila melanogaster*¹². In this species, there is a spaced array of precocious photoreceptors that is strikingly similar to the array of precocious cones in the primate retina. These precocious photoreceptors might secrete a protein that determines the phenotypic commitment of adjacent cells¹². Thus, the spacing of early-maturing cells in these arrays is critical for the development of a mature mosaic because their influence on the phenotypic specification of adjacent cells is locally restricted.

In view of recent evidence that both cell-cell interactions and cell lineage play a part in determining retinal cell types in

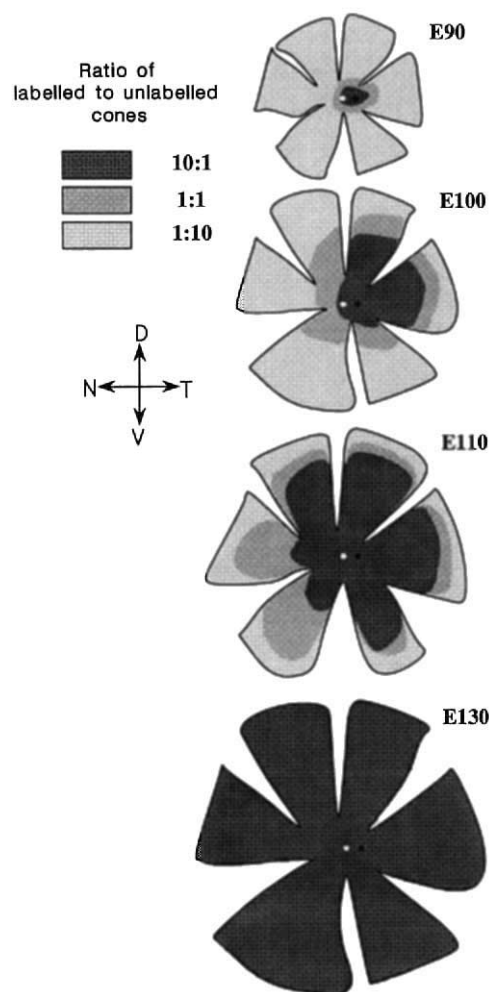


FIG. 3 A schematic diagram of the developmental progression in the ratio of 4942A-labelled to unlabelled cones in the fetal retina at E90, E100, E110 and E130. Because the red/green opsin appears to be expressed at least three weeks in advance of the blue opsin, our developmental analyses were based on the distribution of 4942A-positive cones. Video-enhanced DIC optics in combination with a video-overlay system were used to quantify the density and ratio of labelled and unlabelled cone inner and outer segments. Counts of immunolabelled and unlabelled cones were made across the entire retina and represented as cone densities in topographic surface plots. The array of precocious cones (1 to 10 ratio of labelled to unlabelled cones) is progressively restricted to more peripheral regions of the retina with increasing age until E130, when the adult-like distribution of 4942-labelled cones (10 to 1 ratio of labelled to unlabelled cones) is found across the entire retina. D, dorsal; V, ventral; N, nasal; T, temporal. Optic disc, white dot; fovea, black dot.

vertebrates¹³⁻¹⁸, early interactions between differentiated and undifferentiated cones could provide an opportunity for the specification of the cone mosaic. We propose that early differentiating cells in the primate photoreceptor mosaic, similar to early differentiating neurons in other parts of the nervous system, have intrinsically encoded positional information about region- and species-specific neural organization¹⁹. □

Received 10 January; accepted 12 April 1991.

- Jacobs, G. H. *Comparative Color Vision* (Academic, New York, 1981).
- Wikler, K. C. & Rakic, P. *J. Neuroscience* **10**, 3390-3401 (1990).
- LaVail, M. M., Yasamura, D. & Rakic, P. *Invest. Ophthalmol. Vis. Sci. Suppl.* **24**, 7 (1983).
- Nishimura, Y. & Rakic, P. *J. comp. Neurol.* **241**, 420-434 (1985).
- Lerea, C. L., Bunt-Milam, A. K. & Hurley, J. B. *Neuron* **3**, 367-376 (1989).
- Nathans, J., Thomas, D. & Hogness, D. S. *Science* **232**, 193-202 (1986).
- Curcio, C. A. et al. *Soc. Neurosci. Mtg* **15**, 1206 (1989).
- Wikler, K. C., Williams, R. W. & Rakic, P. *J. comp. Neurol.* **297**, 499-508 (1990).

9. LaVail, M. M., Rapaport, D. H. & Rakic, P. *J. comp. Neurol.* (in the press).
10. Finlay, B. L. & Sengelaub, D. R. *Development of the Vertebrate Retina* (Plenum, New York, 1989).
11. Rapaport, D. H. & Stone, J. *Dev. Brain Res.* **5**, 273-279 (1982).
12. Baker, N. E., Mlodzik, M. & Rubin, G. M. *Science* **250**, 1370-1377 (1990).
13. Reh, T. A. & Kijavini, I. J. *J. Neurosci.* **9**, 4179-4189 (1989).
14. Wetts, R., Serbedzija, G. N. & Fraser, S. E. *Dev. Biol.* **136**, 254-263 (1989).
15. Adler, R. & Hatlee, M. *Science* **243**, 391-393 (1989).
16. Watanabe, T. & Raff, M. C. *Neuron* **2**, 461-467 (1990).
17. Turner, D. L. & Cepko, C. L. *Nature* **328**, 131-136 (1987).
18. Holt, C. E., Bertsch, T. W., Ellis, H. M. & Harris, W. A. *Neuron* **1**, 15-26 (1988).
19. Rakic, P. *Science* **241**, 170-176 (1988).
20. Johnson, L. V. & Hageman, G. S. *Invest. Ophthalmol. Vis. Sci.* **29**, 550-557 (1988).

ACKNOWLEDGEMENTS. We thank C. Lerea and L. Johnson for providing antibodies, and R. Cameron, P. Goldman-Rakic and M. Schwartz for comments on the manuscript. This study was supported by the NIH.

Bcr encodes a GTPase-activating protein for p21^{rac}

Dagmar Diekmann*, Suzanne Brill*, Michelle D. Garrett*, Nicholas Totty†, Justin Hsuan†, Clinton Monfries‡, Christine Hall‡, Louis Lim‡ & Alan Hall*§

* Chester Beatty Beatty Laboratories, Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK

† Ludwig Institute for Cancer Research, 91 Riding House Street, London W1P 8BT, UK

‡ Institute of Neurology, 1 Wakefield Street, London WC1 1PJ, UK

§ To whom correspondence should be addressed

MORE than thirty small guanine nucleotide-binding proteins related to the *ras*-encoded oncoprotein, termed Ras or p21^{ras}, are known¹. They regulate many fundamental processes in all eukaryotic cells, such as growth, vesicle traffic and cytoskeletal organization. GTPase-activating proteins (GAPs) accelerate the intrinsic rate of GTP hydrolysis of Ras-related proteins, leading to down-regulation of the active GTP-bound form². For p21^{ras}, two GAP proteins are known, rasGAP and the neurofibromatosis (NF1) gene product²⁻⁵. There is evidence that rasGAP may also be a target protein for regulation by Ras and be involved in downstream signalling⁶⁻⁸. We have purified a GAP protein for p21^{rho}, which is involved in the regulation of the actin cytoskeleton⁹. Partial sequencing of rhoGAP reveals significant homology with the product of the *bcr* (breakpoint cluster region) gene, the translocation breakpoint in Philadelphia chromosome-positive chronic myeloid leukaemias. We show here that the carboxy-terminal domains of the *bcr*-encoded protein (Bcr) and of a Bcr-related protein, n-chimaerin, are both GAP proteins for the Ras-related GTP-binding protein, p21^{rac}. This result suggests that Bcr could be a target for

regulation by Rac and has important new implications for the role of *bcr* translocations in leukaemia.

The identification of rasGAP as a target for regulation by p21^{ras} has prompted a search for other GAP proteins that might interact with Ras or with other members of the Ras-related superfamily^{1,2,6,7}. We have purified from human spleen extracts a protein of relative molecular mass 29,000 (29K) with GAP activity; it is active on p21^{rho} but not p21^{ras}, R-Ras or Rap proteins^{10,11}. As a preliminary to cloning rhoGAP, we have partially sequenced the protein. A search of the SWISS-PROT protein database (release 15) using a peptide sequence from rhoGAP (shown in Fig. 1) revealed striking similarity to a region present in Bcr, the site on chromosome 22 to which the *c-abl* proto-oncogene is translocated in Philadelphia chromosome-positive chronic myeloid leukaemias and acute lymphoid leukaemias¹²⁻¹⁴. A Bcr-related protein, n-chimaerin, also has significant homology with the rhoGAP peptide¹⁵. Out of 31 amino acids in the rhoGAP peptide, Bcr and n-chimaerin show 50% and 33% identity respectively (Fig. 1).

To test whether Bcr or n-chimaerin act as GAP proteins towards p21^{rho}, we expressed domains of these proteins in *Escherichia coli*. The normal *bcr* gene encodes a protein of 1,271 amino acids and the rhoGAP homology has been localized to the extreme carboxy terminus¹⁴. A stretch of complementary DNA encoding a Bcr C-terminal domain of 401 amino acids was inserted into an *E. coli* expression vector under the control of the tryptophan promoter¹⁶. The cDNA product was fused at its amino terminus to a 10-amino-acid Myc epitope tag to facilitate identification of recombinant protein. Figure 2a (lane 1) is a western blot showing expression of the Bcr C-terminal domain in a bacterial extract. The n-chimaerin was expressed as a glutathione-S-transferase (GST) fusion protein using either the complete sequence or the C-terminal domain of n-chimaerin¹⁷. In this case, recombinant protein was purified from bacterial extracts using glutathione-Sepharose affinity chromatography. Figure 2b shows a Coomassie blue-stained SDS-polyacrylamide gel of purified GST control protein (lane 1) and of the GST/C-terminal n-chimaerin fusion protein (lane 2). The C-terminal n-chimaerin domain could be separated from the GST fusion protein by proteolytic cleavage with thrombin (Fig. 2b, lane 3).

E. coli extracts containing Bcr and purified C-terminal n-chimaerin were incubated with recombinant p21^{rho} that had been preloaded with [γ -³²P]GTP and the effect on the rate of GTP hydrolysis investigated¹⁰. Figure 2c shows that, in contrast to rhoGAP, neither Bcr nor n-chimaerin could stimulate the GTPase activity of p21^{rho}.

Four other small GTP-binding proteins have relatively high (>55%) sequence homology with p21^{rho} (refs 18-20). To see whether Bcr or n-chimaerin could act as GAP proteins for p21^{rho} relatives, we expressed the cDNA for one of these, human *rac1*, in an *E. coli* expression vector¹⁸. In addition, we expressed *rac*

FIG. 1 A comparison of a rhoGAP peptide sequence with Bcr and n-chimaerin. Dots represent sequence identity with rhoGAP amino acids.

METHODS. rhoGAP was purified by >2,000 fold from a human spleen extract using FPLC chromatography¹¹. The protein (2 µg) was electroeluted from an SDS-polyacrylamide gel and digested with lysyl endopeptidase (protease I from *Achromobacter lyticus* M497-1; 0.4 µg ml⁻¹; ref. 28). Peptides were separated by anion-exchange chromatography and reverse-phase liquid chromatography, and sequenced on an Applied Biosystems Model 477A Pulse Liquid Sequencer. Because lysyl endopeptidase cleaves C-terminal to lysine residues, a lysine is included at the start of the sequence. The Bcr sequence corresponds to amino acids 1,197 to 1,226 of the 1,271-amino-acid human protein, and the n-chimaerin sequence corresponds to amino acids 252-281 of the 299-amino-acid human proteins^{14,15}. Homology between Bcr and n-chimaerin extends over the C-terminal 171 amino acids of both proteins¹⁵.

GAP ^{rho}	K M T N T N L A V V F G P N L L W A K D A A I T L K A I N P I
Bcr	· · S L H · · · T · · · · T · · R P S E K E S K · P · · · S
n-Chimaerin	L · N A E · · · G I · · · · T · M R S P E L D A M A A L · D ·

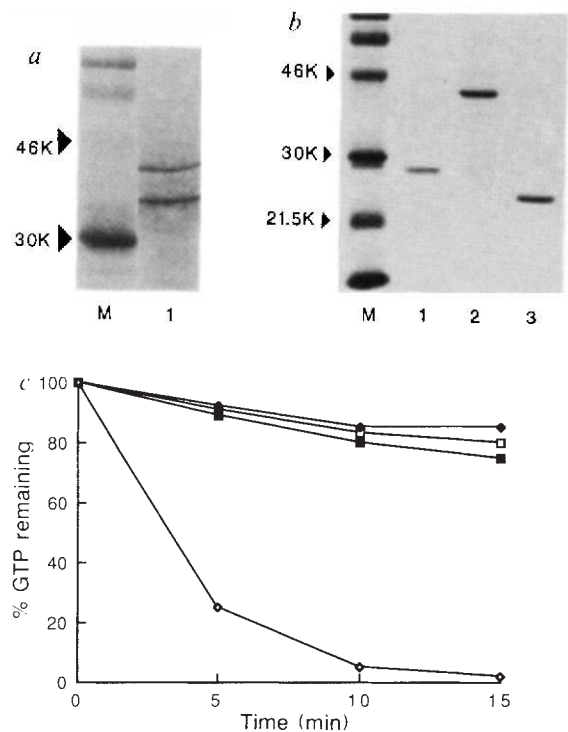
FIG. 2 Expression of Bcr and n-chimaerin C-terminal domains in *E. coli* and their effects on p21^{rho} GTPase activity. **a**, Western blot of an *E. coli* extract expressing epitope-tagged Bcr carboxy terminus, developed using an anti-tag antibody showing a band of the predicted relative molecular mass (~40 K) and a smaller degradation product (lane 1); lane M shows migration of *M_r* standards. **b**, Coomassie blue-stained SDS-polyacrylamide gel of affinity-purified GST/n-chimaerin fusion proteins; lane 1, GST control; lane 2, GST/C-terminal n-chimaerin fusion; lane 3, as lane 2, but the fusion protein is cleaved with thrombin and the C-terminal domain purified. Lane M shows migration of marker proteins. **c**, Intrinsic GTPase activity of p21^{rho} (□) and the effect of adding rhoGAP (◇), C-terminal Bcr (◆) and C-terminal n-chimaerin (■).

METHODS. Human Bcr cDNA encoding Bcr amino acids 871–1,271 was ligated in-frame to a double-stranded oligonucleotide encoding the Myc epitope sequence: MetGluGlnLysLeuileSerGluGluAspLeu. The epitope-tagged cDNA was inserted into a pAT153-based plasmid under the control of a tryptophan promoter as described¹⁶. *E. coli* cells (500 ml) containing this C-terminal-Bcr expression plasmid were grown for 4 h in the presence of indole-acrylic acid (50 µg ml⁻¹) to induce expression from the tryptophan promoter. Cells were pelleted and resuspended in 3 ml of 20 mM Tris, pH 7.6, 0.1 mM DTT, 10 mM NaCl, 1 mM MgCl₂, and the suspension sonicated. After centrifugation to remove cell debris, the supernatant was added directly to the p21^{rho} GTPase assay (see below). Control extracts were made from cells harbouring a similar plasmid but with no *bcr* insert and used in parallel. Expression of the C-terminal Bcr recombinant protein was confirmed using western blots and the mouse monoclonal antibody Myc1/9E10, which specifically recognizes the Myc epitope tag. Human n-chimaerin was introduced into glutathione-S-transferase expression vectors (Pharmacia). For C-terminal n-chimaerin, a restriction fragment encoding amino acids 108–299 was inserted into pGEX-2T to generate a GST/C-terminal n-chimaerin fusion protein¹⁷. *E. coli* cells harbouring these plasmids were grown to an absorbance at 500 nm of 1.0, and expression induced with isopropyl-β-D-thiogalactoside (1 mM) for 1 h. Pelleted cells were lysed by sonication. After addition of 1% Triton X100, GST control or GST/n-chimaerin C-terminal protein-containing extracts were passed over a glutathione-Sepharose 4B column (Pharmacia) and the proteins eluted with 5 mM glutathione, 50 mM Tris-HCl, pH 8.0 (ref. 17). Purified C-terminal n-chimaerin was derived from the GST fusion by thrombin cleavage²⁹. To measure GTPase activity, recombinant p21^{rho} was preincubated with [γ-³²P]GTP (Amersham; 30 Ci mmol⁻¹) in low Mg²⁺ as described^{10,21}. This GTP-loading reaction was stopped with excess free Mg²⁺ and the GTPase activity determined at 30 °C by measuring the loss of protein-bound radioactivity using a nitrocellulose filter-binding assay. The source of rhoGAP was a 2,000-fold purified fraction from human spleen extracts. An *E. coli* extract containing the Bcr C-terminal protein (10 µl) or purified n-chimaerin C-terminal protein (3 µg) was added.

with a glycine-to-valine mutation at codon 12 (p21^{rac}_{V12}; equivalent to the oncogenic mutation in p21^{ras}). Figure 3a shows a western blot of partially purified p21^{rhoA} (lane 1), normal p21^{rac1} (lane 2), and p21^{rac}_{V12} (lane 3), which was developed using an antipeptide antibody raised against a sequence common to both p21^{rac} and p21^{rho} (ref. 21). In Fig. 3b it can be seen that the Val 12 mutation reduces the intrinsic GTPase activity of p21^{rac} by 10-fold. Figure 3b also shows that GAP^{rho} can stimulate the intrinsic GTPase activity of normal (Gly 12)p21^{rac}, although it was about ten times less active than with p21^{rho}. GAP^{rho} did not stimulate the GTPase activity of p21^{rac}_{V12}.

We next looked at the effect of the Bcr C-terminal domain on p21^{rac} GTPase rates. Figure 3c shows that an *E. coli* extract containing C-terminal Bcr dramatically stimulates the GTPase activity of p21^{rac}, whereas a control *E. coli* extract has no effect. Bcr was unable to stimulate the GTPase activity of p21^{rac}_{V12} and it had no effect on the guanine nucleotide-exchange rates of either the normal or mutant p21^{rac} (data not shown). Figure 3c shows that the C-terminal domain of n-chimaerin is also a GAP protein for p21^{rac}. A fusion protein of GST and full-length n-chimaerin was found to be active, but GST itself had no effect (data not shown). n-Chimaerin had no effect on the GTPase activity of the mutant p21^{rac}, or on the nucleotide-exchange rates of either p21^{rac}. Neither Bcr nor n-chimaerin had any effect on the GTPase activities of p21^{ras} proteins.

We conclude that GAP^{rho}, Bcr and n-chimaerin belong to a family of GAP proteins specific for the Rho/Rac family of small



GTP-binding proteins. GAP^{rho} is active on p21^{rho} and p21^{rac}, whereas Bcr and chimaerin are active only on p21^{rac1}, although they may still interact with p21^{rho}. The Rac2 protein has >90% amino-acid identity with that from Rac1 and we expect Bcr and chimaerin to be equally active on p21^{rac2} (ref. 18). Two other members of the Rho subgroup, G25K and TC10, may also be substrates for Bcr-related GAP proteins^{19,20}. Recently another protein with homology to Bcr has been identified: the 85K subunit of phosphatidylinositol 3-kinase²². It remains to be seen whether it belongs to this family of GAP proteins.

Nothing is known about the function of the Rac proteins; most cells express either *rac1* or *rac2* messenger RNA or both, and interestingly, levels are higher in cells of the myeloid lineage¹⁸. Our results show that Bcr regulates the activity of p21^{rac} and suggest that it might also be the target for p21^{rac} regulation. The functions of Bcr and chimaerin are unknown; the 130K protein encoded by *bcr* is expressed in the cytoplasm of most cell types^{23,24}. Translocations involving *bcr* and *c-abl* have been found in chronic myeloid and acute lymphoid leukaemias and result in a Bcr/Abl fusion protein with deregulated tyrosine kinase activity^{25,26}. In this fusion product, the N-terminal domain of Bcr specifically activates the Abl kinase and the p21^{rac} GAP domain of Bcr reported here is absent²⁷. But the translocation is a reciprocal event, raising the possibility that a RacGAP domain fused to the extreme N-terminal amino acids of *c-abl* protein is expressed in Philadelphia chromosome-positive leukaemias. In addition, the normal *bcr* allele seems to

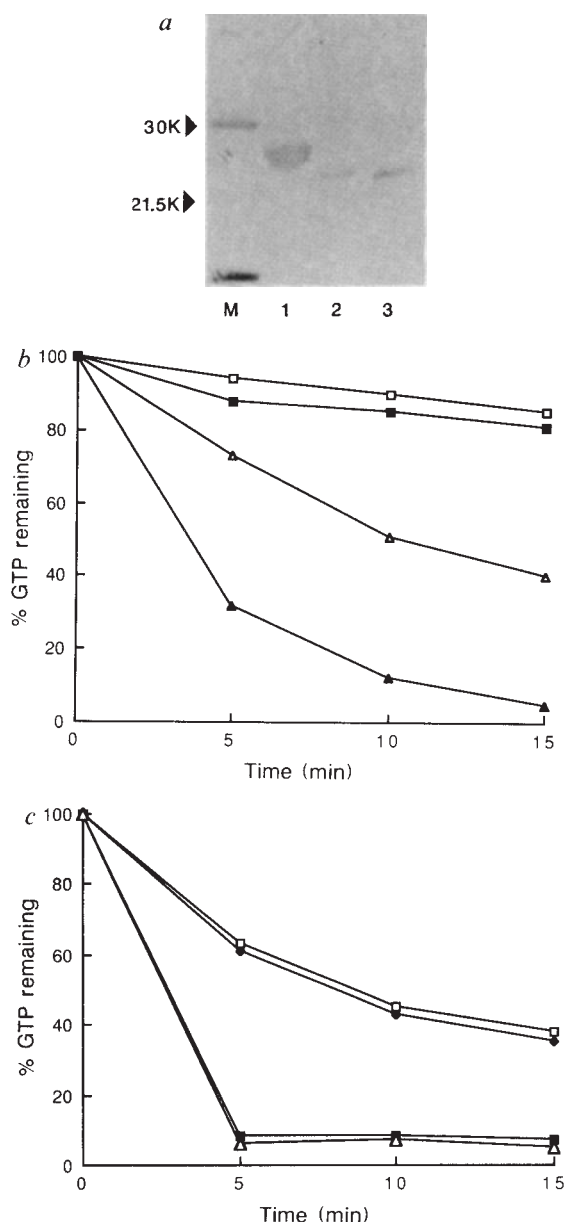


FIG. 3 Expression of recombinant p21^{rac} and the effect of rhoGAP. *a*, Western blot of partially purified normal p21^{rhoA} (lane 1), normal p21^{rac} (lane 2) and mutant p21^{rac} (Gly12 → Val) (lane 3). Lane M, M, markers. *b*, GTPase activity of normal p21^{rac} (triangles) and p21^{rac_v12} (squares) with (filled symbols) or without (open symbols) rhoGAP. *c*, The GTPase activity of normal Rac (□) and the effect of adding Bcr C-terminal protein (△), C-terminal n-chimaerin (■), or a control *E. coli* extract (●).

METHODS. Human *rac1* cDNA was obtained as a polymerase chain reaction fragment from a human HT1080 fibrosarcoma cDNA library¹⁸. Using oligonucleotide site-directed mutagenesis (Amersham), a *Hind*III site was made directly upstream of the initiation codon and a Gly → Val mutation created at codon 12. All cDNAs were sequenced and then introduced into a plasmid under the control of the tryptophan promoter¹⁶. Normal and mutant p21^{rac} were partially purified from this *E. coli* expression system as described for p21^{rhoA} and p21^{rac} (refs 10, 21). Protein was visualized on western blots using a rabbit polyclonal antibody raised against a peptide corresponding to amino acids 65–75 of p21^{rhoA}; this sequence is conserved in p21^{rac}. The GTPase activity of p21^{rac} was determined as described for p21^{rhoA} in Fig. 2 legend, but the assays were at 20 °C. The results in *c* were obtained using a control *E. coli* extract (10 µl), an *E. coli* extract containing C-terminal Bcr (1 µl) and purified C-terminal n-chimaerin (1 µg). Purified GST control protein (10 µg) had no effect. To measure nucleotide exchange rates, p21^{rac} was preincubated with [³H]GTP (Amersham; 7 Ci mmol⁻¹) instead of [³²P]GTP, and the loss of nucleotide was measured in the presence of excess unlabelled GTP by the nitrocellulose filter-binding assay²¹.

be maintained in these leukaemias, although it is unclear whether a normal protein is expressed²³. In any event, the role of *bcr* in chronic myeloid and acute lymphoid leukaemias now needs to be re-examined: in particular, the possible inactivation or constitutive activation of its racGAP activity would be expected to have profound biological effects that might contribute to the development of leukaemia. □

Received 25 March; accepted 23 April 1991.

- Hall, A. *Science* **249**, 636–640 (1990).
- Trahey, M. & McCormick, F. *Science* **238**, 542–545 (1987).
- Martin, G. A. *et al.* *Cell* **63**, 843–849 (1990).
- Ballester, R. *et al.* *Cell* **63**, 851–859 (1990).
- Xu, G. *et al.* *Cell* **63**, 835–841 (1990).
- Calés, C., Hancock, J. F., Marshall, C. J. & Hall, A. *Nature* **332**, 548–551 (1988).
- Adari, H., Lowy, D. R., Wilmsen, B. M., Der, C. J. & McCormick, F. *Science* **240**, 518–521 (1988).
- Yatani, A. *et al.* *Cell* **61**, 769–776 (1990).
- Paterson, H. F. *et al.* *J. Cell Biol.* **111**, 1001–1007 (1990).
- Garrett, M. D., Self, A. J., Van Oers, C. & Hall, A. *J. biol. Chem.* **264**, 10–13 (1989).
- Garrett, M. D., Major, G. N., Totty, N. & Hall, A. *Biochem. J.* (in the press).
- Heisterkamp, N., Stam, K., Groffen, J., Klein, A. & Grosveld, G. *Nature* **315**, 758–761 (1985).
- Hariharan, I. K. & Adams, J. M. *EMBO J.* **6**, 115–119 (1987).
- Lifshitz, B. *et al.* *Oncogene* **2**, 113–117 (1988).
- Hall, C. *et al.* *J. molec. Biol.* **211**, 11–16 (1990).
- Trahey, M. *et al.* *Molec. cell. Biol.* **7**, 541–544 (1987).
- Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).
- Didsbury, J., Weber, R. F., Bokoch, G. M., Evans, T. & Snyderman, R. *J. biol. Chem.* **264**, 16378–16382 (1989).
- Munemitsu, S. *et al.* *Molec. cell. Biol.* **10**, 5977–5982 (1990).
- Drivas, G. T. *et al.* *Molec. cell. Biol.* **10**, 1793–1798 (1990).
- Hall, A. & Self, A. J. *J. biol. Chem.* **261**, 10963–10965 (1986).
- Otsu, M. *et al.* *Cell* **65**, 1–20 (1991).
- Amson, R. B., Marcelle, C. & Teerman, A. *Oncogene* **4**, 243–247 (1989).
- Li, W., Drazan, O., Klotzer, W., Gale, R. P. & Arlinghaus, R. B. *Oncogene* **4**, 127–138 (1989).
- Konopka, J. B., Watanabe, S. M. & Witte, O. N. *Cell* **37**, 1035–1041 (1984).
- McWhirter, J. R. & Wang, W. Y. *J. Molec. cell. Biol.* **11**, 1553–1565 (1991).
- Muller, A. J. *et al.* *Molec. cell. Biol.* **11**, 1785–1792 (1991).
- Kawasaki, H. & Suzuki, K. *Analyt. Biochem.* **186**, 264–268 (1990).
- Gearing, D. P. *et al.* *Biotechnology* **7**, 1157–1160 (1989).

ACKNOWLEDGEMENTS. We thank G. Evan for monoclonal antibody 9E10, D. Jadayel for the *rac1* PCR fragment, P. Allen for a human *bcr* cDNA clone, M. Teo for help with n-chimaerin preparation, and C. J. Marshall for advice. The work at the Chester Beatty Laboratories is funded by the Cancer Research Campaign and the Medical Research Council. L.L. was supported by the Brain Research Trust and the Glaxo-Singapore Research Fund.

Peptide-induced conformational change of the class I heavy chain

Tim Elliott, Vincenzo Cerundolo, John Elvin & Alain Townsend

Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

THERE is evidence that peptide ligands take part in the assembly of class I molecules^{1–8}. In particular, addition of peptides to extracts of the mutant cells RMA-S and .174/T2, in which stable assembly of class I does not occur^{9–11}, results in a conformational change in the class I heavy chain and stable association of the heavy chain with β_2 -microglobulin (β_2m) (refs 1–3). Thus specific peptides may stabilize or induce a conformational change in the class I heavy chain that results in a rise in the binding affinity of the heavy chain for β_2m (Fig. 1*a*). Here we show that peptides have two cooperative roles in class I assembly. Specific short peptides (9–10 amino acids) can induce folding of the heavy chain in the absence of β_2m . Both short (nine amino acids) and longer sequences (15 amino acids) can stabilize preformed low-affinity complexes of heavy chain and β_2m . To alter the conformation of free heavy chains, the peptides must be exactly the correct size, and they are found to correspond to the sequences isolated from infected cells¹². This property may therefore be the basis for selection of epitopes presented *in vivo*.

A dilute lysate of RMA-S cells was prepared, in which heavy and light chains are predominantly dissociated. Peptides were then added, and the change in conformation of the heavy chain and association with $\beta 2m$ followed. Over 20 h at 4 °C a conformational change in the heavy chain (Fig. 1*b, d*), and reassembly with $\beta 2m$ (Fig. 1*c, e*) occurs after addition of nucleoprotein sequences of between 9 and 15 amino acids known to bind to the polypeptide encoded by *H-2D^b* (D^b) (ref. 2). This demonstrates that assembly is reversible, and that peptides can drive the formation of new complexes from dissociated heavy and light chains.

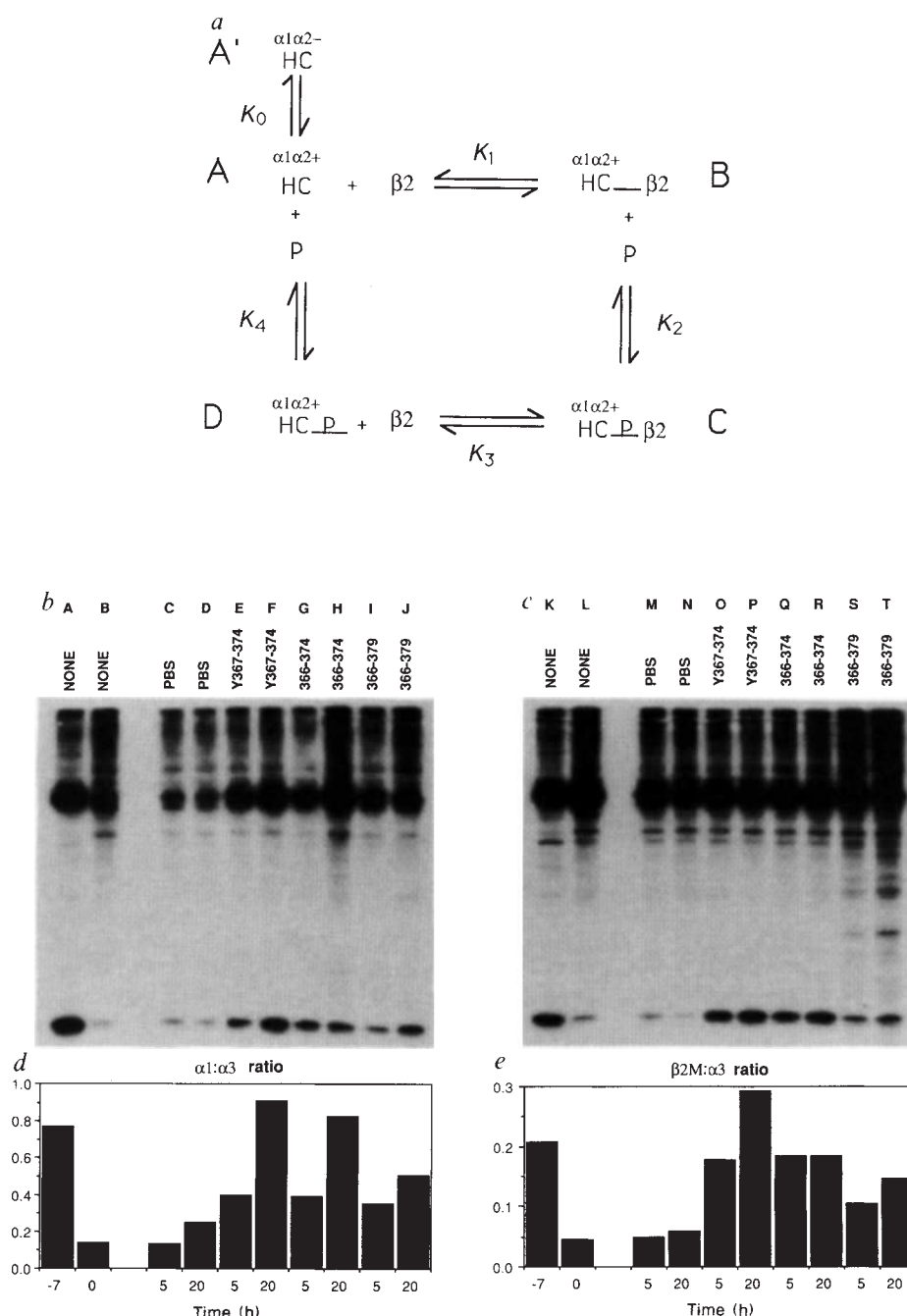
To investigate the effect of peptides on free heavy chains (represented by $A' \rightarrow A \rightarrow D$ in Fig. 1), we added various peptides to a detergent extract of R1E/ D^b (refs 13, 14) cells, which express D^b heavy chains in the absence of $\beta 2m$, and looked for a change in the conformation of the heavy chain. Addition of the 9-amino-

acid peptides 366–374 and a derivative with Ala $\rightarrow$ Tyr at position 366 (Y367–374, single-letter amino-acid notation) gave the characteristic conformational change in D^b detected by an $\alpha 1$ domain-specific antibody in >50% of the heavy chains (Fig. 2). This effect was seen with nucleoprotein sequences derived from both A/PR/8/34 (366–374/34 in Fig. 2) and A/NT/60/68 influenza isolates, which differ by substitutions at positions 372 (Glu $\rightarrow$ Asp) and 373 (Thr $\rightarrow$ Ala) (ref. 15). The study of the effect of these peptides on the conformation of heavy chains has been extended with the D^b -specific antibody 27-11-13S (ref. 16) (data not shown), which is specific for the $\alpha 2$ domain (J. A. Bluestone, personal communication).

We next investigated the length of peptide required for this effect. R1E/ D^b lysates were incubated with peptides whose N and C termini were lengthened or shortened by one amino acid (Fig. 2*b*). Extending or shortening the peptide at the C terminus

FIG. 1 *a*, Schematic diagram of class I assembly in which unfolded heavy chains ($\alpha 1\alpha 2$ -HC) at A' are in equilibrium (K_0) with folded heavy chains ($\alpha 1\alpha 2$ +HC) at A . The $\alpha 1\alpha 2$ + conformation is defined by the binding of conformation-sensitive monoclonal antibodies to the $\alpha 1$ or $\alpha 2$ domains. Folded heavy chains can associate with peptide ($A \rightarrow D$) or $\beta 2m$ ($A \rightarrow B$) before binding the third component of the fully assembled class I molecule (C). Formation of this trimolecular complex at C is favoured if $K_3 > K_1$ and $K_2 > K_4$. *b, c, d, e*, The effect of peptides on reassembly of D^b from free heavy chains and $\beta 2m$.

METHODS. Labelled RMA-S cells were lysed at a concentration of 2×10^7 cells ml^{-1} in 1-ml aliquots. Immediately on lysis, one aliquot was split and heavy chains were precipitated with B22.249 or 28-14-8s as described² to evaluate conformation and assembly of molecules immediately after release from the RMA-S cells (lanes A and K). The remaining aliquots were incubated for a further 7 h at 4 °C, then one aliquot was split and immunoprecipitated with the two antibodies, to demonstrate loss of conformation (lane B) and dissociation (lane L). Peptides (to 75 μM) or PBS medium were then added to the remaining aliquots as indicated, and the mixtures incubated at 4 °C. Heavy chains were precipitated after 5 or 20 h with either B22.249 to detect conformation (lanes C–J), or 28-14-8s to measure $\beta 2m$ association (lanes M–T). The precipitates were washed and analysed by SDS-PAGE and autoradiography as previously described^{13,14}, and the radioactive bands were quantified as follows. After analysing immunoprecipitates by SDS-PAGE and autoradiography, the heavy-chain and $\beta 2m$ bands were excised from the dried gels with the aid of the autoradiograph, and counted on a Pharmacia-Wallac Betaplate TM flat bed scintillation counter. Acquisition of the $\alpha 1\alpha 2$ + conformation is expressed as the ratio of c.p.m. for heavy chains precipitated with the $\alpha 1$ -specific antibody to c.p.m. for heavy chains precipitated with the $\alpha 3$ -specific antibodies ($\alpha 1:\alpha 3$ ratio in *d*). Association of $\beta 2m$ and heavy chain is expressed as the ratio of c.p.m. for $\beta 2m$ to c.p.m. for heavy chains from immunoprecipitations with the $\alpha 3$ -specific antibody ($\beta 2m:\alpha 3$ ratio in *e*). All peptide concentrations were determined using the BCA colorimetric assay (Pierce) unless otherwise stated.



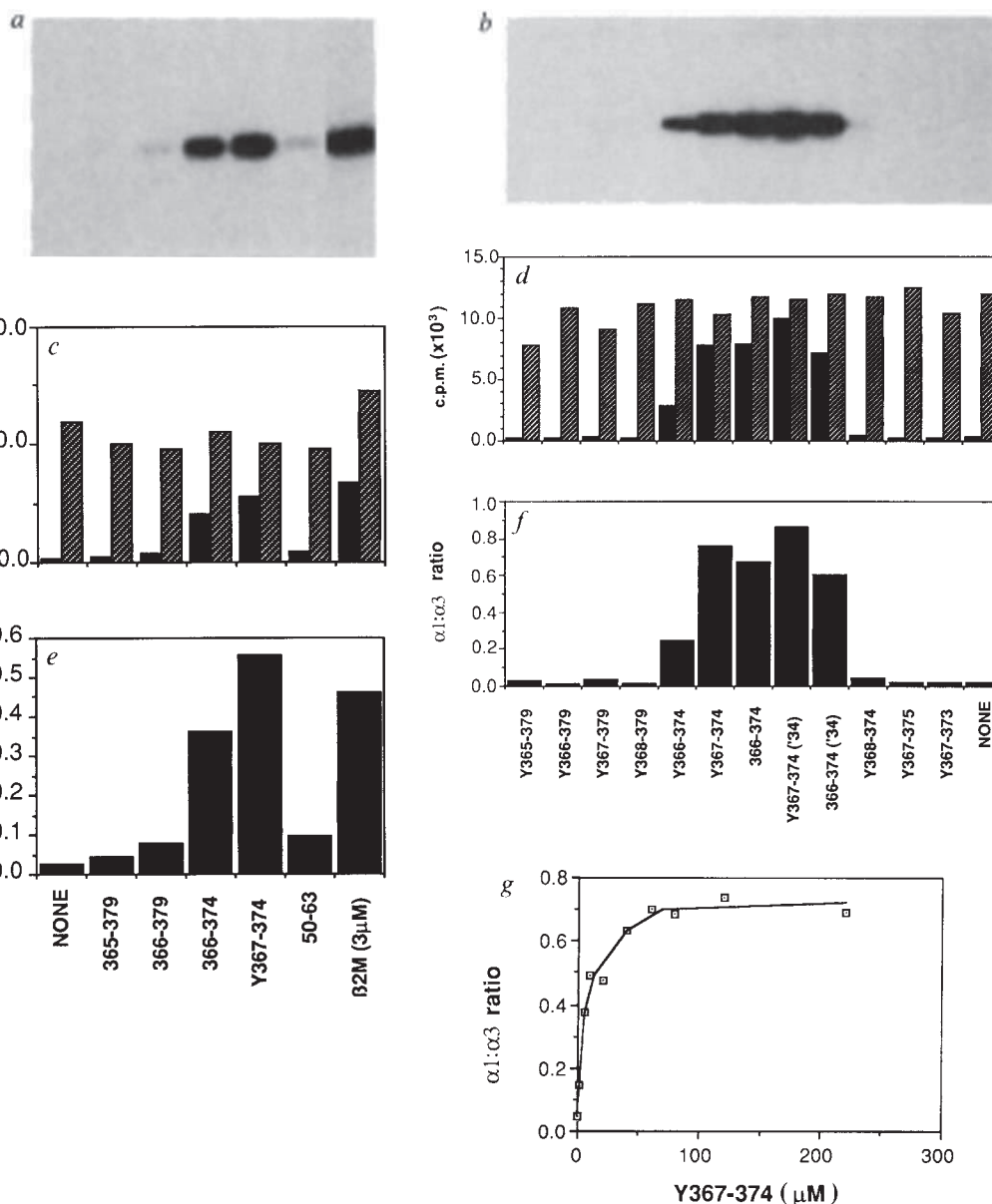


FIG. 2 The effect of peptides on free D^b heavy chains in R1E/D^b lysates (which lack β2m). *a, b*, Lysates of 1×10^7 R1E/D^b cells ml^{-1} were incubated overnight at 4 °C with peptides (75 μM) or human β2m (3 μM) before immunoprecipitating heavy chains with either B22.249 (to assess conformation shown in *a* and *b*) or 28-14-8s (to precipitate total heavy chains; not shown). *c, d*, The heavy chain bands precipitated with B22.249 (solid bars) and 28-14-8s (hatched bars) were excised from the gels and counted as described for Fig. 1. *e, f*, Stabilization of the α1α2+ heavy chain conformation is expressed as the ratio of α1α2+ to α1α2- heavy chain counts (α1:α3 ratio). *g*, Incubations were done with increasing concentrations of peptide NP Y367-374 (A/NT/60/68) before precipitating heavy chains with B22.249 or 28-14-8s. The α1:α3 ratio is plotted against the concentration of peptide added. The half-maximal effect occurred at a concentration of ~4 μM peptide.

completely abolished the effect on the free class I heavy chain. Extending the N terminus by one amino acid (Y366-374) reduced the effect, and the octamer Y368-374, which is shorter than the optimum sequence by one residue, was completely ineffective. The longer sequences Y365-379 and Y366-379, both of which are recognized by cytotoxic T-lymphocytes^{17,18} and can induce some reassembly of class I *in vitro* (Fig. 1), either had no effect (Y365-379), or had a minimal effect (Y366-379). The effect of Y367-374 (Fig. 2g) and 366-374 (not shown) is dose-dependent and saturable, with a half-maximal effect at ~4 μM at 4 °C.

We then looked for a stabilizing effect of peptides on preformed complexes of heavy chain and β2m. Assembly of class I in dilute lysates of RMA-S cells can occur without addition of peptide, by raising the concentration of free human β2m (hβ2m) (ref. 2, and Fig. 2a), or murine β2m (data not shown) in equilibrium with the heavy chain. This is consistent with a low-affinity interaction between free heavy chains and β2m in the absence of peptide represented by A' → A → B in Fig. 1a. These results can also be obtained by raising the concentration both of heavy chains and of β2m, by concentrating a lysate of RMA-S cells (Fig. 3a, b). These complexes formed in the absence of peptide

in concentrated lysates are unstable. After dilution at 4 °C they lose their conformation and dissociate exponentially with a half-life ($t_{1/2}$) of ~3–4 h. But if specific peptides of either 9 or 15 amino acids are added to saturating concentration at the time of dilution, the complexes are stabilized with a resulting $t_{1/2}$ of ~50 h (Fig. 3c–g).

In the pre-Golgi compartment of RMA-S cells, the concentrations of heavy chain and β2m are likely to be greater than in cell lysates of 1.6×10^8 cells ml^{-1} . It is therefore likely that in the endoplasmic reticulum the majority of heavy chains form low-affinity complexes with β2m. Immunoprecipitation immediately after cell lysis (without a preclearing step) detects these complexes (Fig. 1, lanes A and K), which dissociate over ~7 h in dilute lysates at 4 °C (Fig. 3, lanes B and L). Most immunoprecipitation protocols employ an overnight preclearing step in dilute lysates, and therefore detect the majority of RMA-S and 721.174 heavy chains in the dissociated form^{1,2,5,6,11}.

The stabilizing effect of these peptides on heavy chain/β2m complexes is reflected in an increase in the binding of labelled exogenous β2m to heavy chains when peptides are added. For both D^b and A2 the amount of labelled murine or human β2m, respectively, bound at each concentration added was greater in

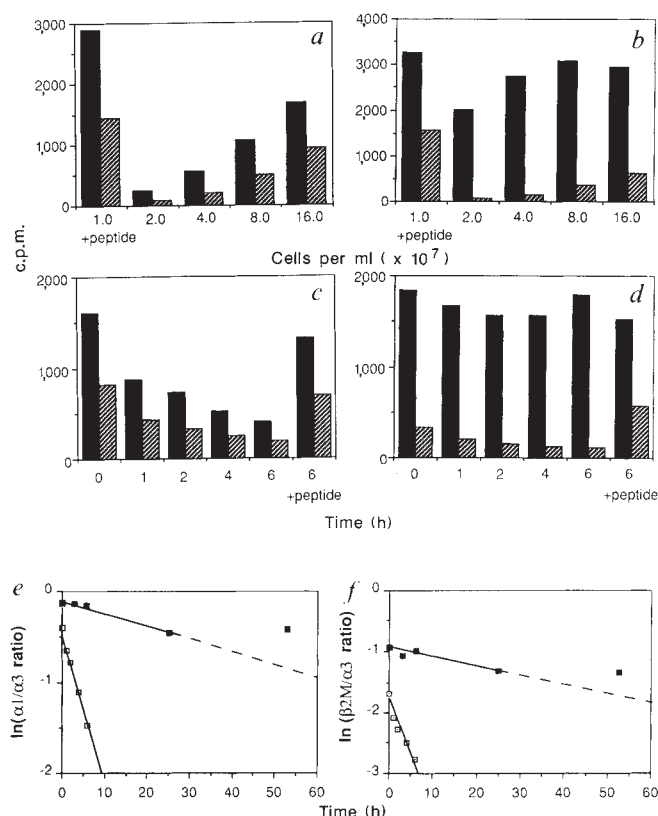
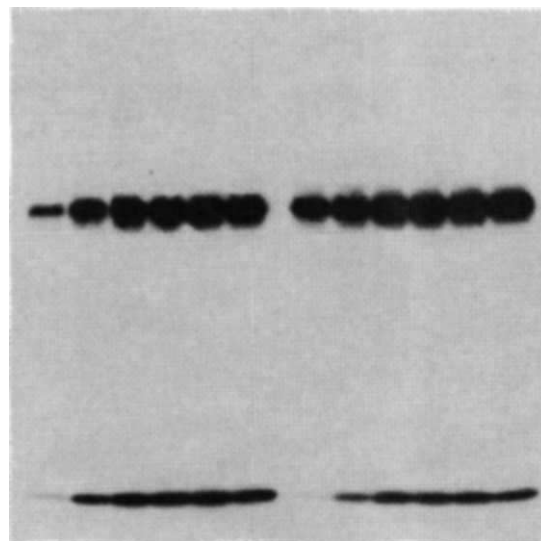


FIG. 3 *a, b*, Increasing the concentration of a lysate of RMA-S increases the level of detectable $\alpha 1\alpha 2+$ heavy chains (*a*) and the degree of saturation of heavy chains with $\beta 2m$ (*b*). *c-f*, Peptide stabilizes heavy chain conformation and slows the rate of dissociation of $\beta 2m$. *g*, Both long and short nucleoprotein (NP) peptides stabilize D^b - $\beta 2m$ complexes.

METHODS. *a, b*, RMA-S cells were lysed at the concentrations indicated in the figure in the absence of peptides or in the presence of 100 μM NP 50-63, and precleared overnight at 4 °C as described for Fig. 1. Heavy chains were then precipitated with either B22.249 to assess conformation (*a*) or 28-14-8s to assess assembly (*b*). Immunoprecipitated bands corresponding to heavy chain (solid bars) and $\beta 2m$ (hatched bars) were cut and counted as described above. *c-f*, RMA-S cells were lysed at 2×10^8 cells ml^{-1} and precleared overnight either in the presence or absence of 100 μM NP 50-63 (determined by weight). For measurements in the absence of peptide, lysates were diluted 20-fold to a final concentration of 1×10^7 cells ml^{-1} at times -6, -4, -2, -1 and 0 h. At time 0, heavy chains were immunoprecipitated with either B22.249 to assess conformation (*c*) or 28-14-8s to assess assembly (*d*). In the presence of peptide, lysates were diluted 20-fold to 1×10^7 cells ml^{-1} at time 0 (final peptide concentration 5 μM). Aliquots were removed for precipitation of heavy chains at time 0, 3, 6, 25 and 53 h.

g A B C D E F G H I J K L



The immunoprecipitates were analysed by SDS-PAGE and autoradiography, and the heavy chains (solid bars) and $\beta 2m$ (hatched bars) were quantitated as described in Fig. 2. Only the 6-h time point is shown in Fig. 4*c* and *d* for complexes formed in the presence of 100 μM NP 50-63. The half-life of the $\alpha 1\alpha 2+$ conformation (*e*) and the heavy chain/ $\beta 2m$ complex (*f*) were derived from the slopes of the lines obtained by plotting the natural log of the $\alpha 1:\alpha 3$ ratios and $\beta 2m:\alpha 3$ ratios against time. In the absence of peptide (open symbols), conformation was lost and $\beta 2m$ dissociated with half-lives of 3.1 h ($R^2=0.93$, $n=5$) and 4.2 h ($R^2=0.93$, $n=5$), respectively. In the presence of peptide (filled symbols), conformation was lost and $\beta 2m$ dissociated with $t_{1/2}$ of 49.0 h ($R^2=0.98$, $n=4$) and 50.0 h ($R^2=0.88$, $n=4$), respectively. *g*, RMA-S cells were ^{35}S -labelled, lysed at 1.65×10^8 cells ml^{-1} and precleared overnight at 4 °C. Lysates were then diluted to 1×10^7 cell equivalents per ml in lysis buffer alone (A and G) or lysis buffer containing 75 μM NP 365-379 (B and H), 75 μM 366-379 (C and I), 75 μM 366-374 (D and J), 38 μM Y367-374 (E and K) or 26 μM 50-63 (F and L). Heavy chains were precipitated with B22.249 to assess conformation (lanes A to F), or 28-14-8s to assess assembly (lanes G to L) after a 7-h incubation at 4 °C, equivalent to ~two half-lives in the absence of peptide.

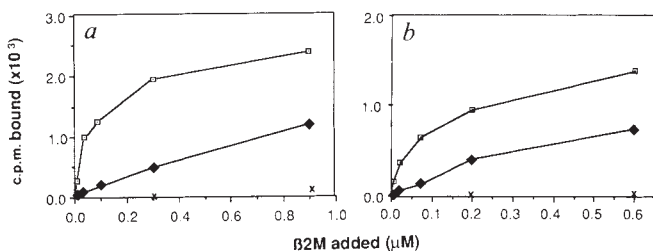


FIG. 4 Binding analysis of radiolabelled murine $\beta 2m$ to D^b (*a*) and human $\beta 2m$ to A2 (*b*) in the presence ($\square$) and absence ($\blacklozenge$) of saturating concentrations of specific peptides.

METHODS. Human $\beta 2$ (Sigma) and mouse $\beta 2m^a$ (prepared by cation exchange from cell culture supernatants of insect cells infected with recombinant baculovirus containing the $\beta 2m$ gene) were radioiodinated using iodobeads (Pierce). Specific activities of labelled $\beta 2m$ were 1.0×10^{16} c.p.m. mol^{-1} for human, and 7.8×10^{14} c.p.m. mol^{-1} for mouse. Aliquots of 1×10^7 RMA-S (*a*) or .174/T2 (*b*) cells were lysed in the presence of increasing concentrations of ^{125}I -labelled $\beta 2m$ either in the absence or presence of 50 μM NP 365-379 (for D^b) or NP 57-78 (K^{62}) (ref. 15) (for HLA-A2), and incubated for 20-22 h at 4 °C while preclearing. Heavy chains were then precipitated with monoclonal antibodies B22.249 (D^b) or BB7.2 (A2), which recognize epitopes on the $\alpha 1$ and $\alpha 2$ domains, respectively, or isotype matched control antibodies (x). Washed immunoprecipitates were counted on a LKB rackgamma, to quantitate coprecipitated $\beta 2m$. When a control peptide (HIVgag p17-3) was included in the incubation, identical curves were obtained to those for incubations performed in the absence of peptide.

lysates to which a stabilizing peptide had been added (Fig. 4a, b).

Our results demonstrate that there are two complementary roles for peptides in class I assembly. Certain short peptides can induce a conformational change in the heavy chain which may provide a high-affinity ligand for $\beta 2m$ ($A' \rightarrow A \rightarrow D \rightarrow C$, Fig. 1a). In addition, they and longer related sequences slow the dissociation of $\beta 2m$ from heavy chains ($A' \rightarrow A \rightarrow B \rightarrow C$, Fig. 1a). As the peptides that influence the conformation of free heavy chains are those that can be isolated from infected cells¹², and form extremely stable peptide class I complexes *in vitro*¹⁹, this property may be the basis for the formation of long-lived peptide class I complexes *in vivo*. □

Received 19 February; accepted 23 April 1991.

1. Townsend, A. *et al. Nature* **340**, 443–448 (1989).
2. Townsend, A. *et al. Cell* **62**, 285–295 (1990).
3. Cerundolo, V. *et al. Nature* **345**, 449–456 (1990).
4. Lie, W. R. *et al. Nature* **344**, 439–441 (1990).
5. Ljunggren, H. G. *et al. Nature* **346**, 476–480 (1990).
6. Schumacher, T. N. M. *et al. Cell* **62**, 563–567 (1990).
7. Kvist, S. & Hamann, U. *Nature* **348**, 446–448 (1990).
8. Bjorkman, P. J. *et al. Nature* **329**, 512–516 (1987).
9. Ljunggren, H. G. & Karre, K. *J. exp. Med.* **162**, 1745–1759 (1985).
10. DeMars, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8183–8187 (1985).
11. Salter, R. D. & Cresswell, P. *EMBO J.* **5**, 943–949 (1986).
12. Rotzschke, O. *et al. Nature* **348**, 252–254 (1990).
13. Allen, H., Fraser, J., Flyer, D., Calvin, S. & Flavell, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7447–7451 (1986).
14. Williams, D. B., Barber, B. H., Flavell, R. A. & Allen, H. *J. Immun.* **142**, 2796–2806 (1989).
15. Huddleston, J. A. & Brownlee, G. G. *Nucleic Acids Res.* **10**, 1029–1038 (1982).
16. Ozato, K. & Sachs, D. H. *J. Immun.* **126**, 317–321 (1980).
17. Townsend, A. R. M. *et al. Cell* **44**, 959–968 (1986).
18. Bastin, J., Rothbard, J., Davey, J., Jones, I. & Townsend, A. *J. exp. Med.* **165**, 1508–1523 (1987).
19. Cerundolo, V. *et al. Eur. J. Immun.* (in the press).

ACKNOWLEDGEMENTS. We thank F. Godeau, P. Kourilsky and P. Robinson for murine $\beta 2m^*$, D. Hayward and P. Thornton-Everson for technical help, D. Wiley, K. Karre, C. Öhlén and H.-G. Ljunggren for discussion, and C. Potter for advice on the use of the Pharmacia-Wallac Betaplate TM flatbed scintillation counter. This work was supported by the MRC, the Wellcome Trust and the Cancer Research Institute (New York).

X-chromosome inactivation may explain the difference in viability of XO humans and mice

Alan Ashworth, Sohaila Rastan*, Robin Lovell-Badge† & Graham Kay*

Chester Beatty Laboratories, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

* Section of Comparative Biology, MRC Clinical Research Centre, Harrow, Middlesex HA1 3UJ, UK

† Laboratory of Eukaryotic Molecular Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

ONLY about 1% of human XO conceptuses survive to birth and these usually have the characteristics of Turner's syndrome, with a complex and variable phenotype including short stature, gonadal dysgenesis and anatomical defects¹. Both the embryonic lethality and Turner's syndrome are thought to be due to monosomy for a gene or genes common to the X and Y chromosomes². These genes would be expected to be expressed in females from both active and inactive X chromosomes to ensure correct dosage of gene product. Two genes with these properties are *ZFX* and *RPS4X*, both of which have been proposed to play a role in Turner's syndrome^{3,4}. In contrast to humans, mice that are XO are viable with no prenatal lethality (P. Burgoyne, personal communication) and are anatomically normal and fertile. We have devised a system to analyse whether specific genes on the mouse X chromosome are inactivated, and demonstrate that both *Zfx* and *Rps4X* undergo

normal X-inactivation in mice. Thus the relative viability of XO mice compared to XO humans may be explained by differences between the two species in the way that dosage compensation of specific genes is achieved.

To study the expression of individual genes from the inactive X chromosome of the mouse, we have produced interspecific F_1 hybrids between laboratory mice (*Mus musculus domesticus*) (dom) carrying the T(X;16)(T16^{dom}) translocation⁵, and *Mus spretus*⁶, a wild mouse species (Fig. 1a). As a result of evolutionary divergence between the species, sequence variants can be readily found in the transcripts of genes (A.A., unpublished results). During early embryonic development of a T16^{dom} heterozygote, cells that have inactivated the T16^{dom} translocation product carrying the X-chromosome inactivation centre are genetically unbalanced and are rapidly selected against, resulting in the normal X chromosome being inactivated in all the remaining cells^{5,7,8}. Hence interspecific F_1 hybrid mice of the genotype T16^{dom}/X^{spe} exhibit total nonrandom inactivation of the X^{spe} chromosome. Expression of a transcript from X^{spe} for a particular gene would indicate that the locus in question escapes inactivation. These animals therefore provide an *in vivo* system for assessing the inactivation status of any X-linked gene for which it is possible to distinguish between the *spretus* (spe) and *domesticus* (dom) transcripts. Figure 1 demonstrates that two genes known to be X-inactivated, *Otc* and *Ar*, are expressed only from the T16^{dom} chromosome, confirming that the system can be used to assay for X-inactivation. In comparison, control interspecific X^{dom}/X^{spe} females express both dom and spe transcripts for *Otc* and *Ar* (Fig. 1), showing that in these mice either X chromosome may be inactivated. The relative expression of the *Otc* alleles as well as the strength of the *Ta* phenotype on a *spretus* background (not shown) suggests that the X^{spe} carries a strong *Xce* allele^{9–11}, that is, the X^{spe} chromosome is more likely to remain active than the X^{dom}.

In humans, few genes outside the pseudoautosomal region are transcribed from both active and inactive X chromosomes^{4,12,13}. One of these genes, *ZFX*, has a closely related homologue on the Y chromosome, *ZFY*^{14,15}. The function of the *ZFY* and *ZFX* genes is unknown although they encode proteins having motifs characteristic of transcription factors¹⁶. Both *ZFY* and *ZFX* appear to be ubiquitously expressed in both male and female human cells, and the extreme similarity of the proteins that they encode has led to the suggestion that they can function interchangeably¹². The lack of X inactivation of the *ZFX* gene would mean that male and female cells have equivalent doses of *ZFX* and *ZFY* protein. The pattern of expression of this gene family is different in mice. The two Y-chromosome homologues, *Zfy-1* and *Zfy-2*, have a very restricted pattern of expression¹⁷ and most somatic murine cells express only *Zfx*^{18,19}. We therefore determined whether *Zfx* undergoes X-inactivation in mice. Figure 2 shows that T16^{dom}/X^{spe} mice express only the X^{dom} allele of *Zfx*. This shows that in mice, unlike humans, *Zfx* undergoes X-inactivation. Thus, XO, XX and XY animals express equivalent amounts of *Zfx* and *Zfy* gene product once X inactivation has taken place. Although the pattern of expression of *ZFY* and *ZFX* is different in mouse and man, equivalent dosage between the sexes is maintained, albeit by a different mechanism. *Zfy-1* is expressed in cell types that in the female would not show X-inactivation, such as embryonic stem cells and germ cells in the fetal gonad¹⁷.

Ribosomal protein S4 (RPS4) in humans is encoded by two very similar genes, one on the X chromosome (*RPS4X*) and one on the Y (*RPS4Y*)⁴. Unlike other genes on Xq, *RPS4X* escapes X inactivation and haploinsufficiency of this gene may have a role in Turner's syndrome⁴. Using the T16^{dom}/X^{spe} system we found that T16^{dom}/X^{spe} mice only express the X^{dom} allele of *Rps4X* (Fig. 3), that is, in mice, unlike humans, *Rps4X* undergoes normal X inactivation (Fig. 3). Furthermore, mice seem to lack a Y-linked homologue of *Rps4X* as determined

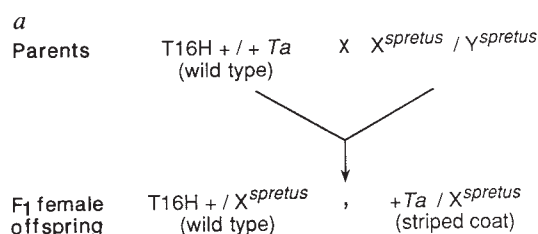


FIG. 1 Inactivation of X-chromosomal genes *Otc* and *Ar* is nonrandom in T16^{dom}/ X^{spe} mice. **a**, Generation of T16^{dom}/ X^{spe} mice. Female offspring carrying the T16^{dom} translocation display a wild-type coat and offspring inheriting the intact *Mus m. domesticus* X chromosome display a characteristic striped coat, due to the presence of the mutant *Ta* gene marker. **b**, Assay of gene expression. RNA prepared from the livers of X^{dom}/X^{dom} , X^{spe}/X^{spe} , X^{dom}/X^{spe} and T16^{dom}/ X^{spe} mice was transcribed into cDNA. *Otc* and *Ar* cDNA was amplified using gene-specific primers and the PCR products directly sequenced. *Otc*^{spe} differs from *Otc*^{dom} in the 3' untranslated region of the gene by the absence of 6 base pairs (bp) resulting in the shift in position of the doublet indicated by the arrowheads. In the control interspecific F₁ animal (X^{dom}/X^{spe}) both alleles are expressed. Interspecific F₁ T16^{dom}/ X^{spe} animals express only the *Otc*^{dom} allele, confirming that *Otc* is subject to X-inactivation²⁰. *Ar*^{dom} and *Ar*^{spe} alleles differ by a sequence variant at the position indicated by the arrowhead. The control interspecific F₁ animal (X^{dom}/X^{spe}) expresses both alleles, yet the T16^{dom}/ X^{spe} animal expresses only the *Ar*^{dom} gene, confirming that *Ar* undergoes X-inactivation²¹.

METHODS. **a**, Oocytes from T16^{dom} +/+ Ta female mice were fertilized *in vitro* with sperm from a *Mus spretus* male by a modification of standard procedures (G.K. *et al.*, manuscript in preparation) and transferred to the oviducts of pseudopregnant recipients. The two types of F₁ female offspring thus generated were distinguished by coat phenotype. The presence of X^{dom} and X^{spe} was verified in the various mice by Southern blot analysis with probes specific for the X chromosome detecting genomic restriction-fragment length variants between *Mus m. domesticus* and *Mus spretus* (data not shown). **b**, RNA was prepared from the livers of mice using the guanidinium thiocyanate/phenol method²² or using a FastTrack kit (Invitrogen). Complementary DNA was synthesized with oligo (dT) as primer, amplified using PCR and sequenced directly as described (ref. 21, and A. S. Goldsborough and A.A., manuscript in preparation). *Otc* was amplified with primers (TGCTAATTGAGGATCCTG and TTTCCACATAAAGTGACTT) corresponding to nucleotides 8–28 and 1,178–1,158 of the previously determined sequence²³. *Ar* primers (TACTTCCCACCCAGAAGACCTG and GCAAATACCTACAGTCCCATCC) were synthesized based on regions minimally degenerate between rat and human *Ar* genes and correspond to nucleotides 1,635–1,657 and 2,224–2,203 of the published rat sequence²⁴. Primers were designed to hybridize to separate exons of *Otc* and *Ar* to eliminate the possibility of amplification of contaminating genomic DNA. The *Otc* sequence shown is the coding strand and is complementary to that shown on the sequencing gel. Sequence variants between *Mus m. domesticus* and *Mus spretus* were identified by DNA sequencing.

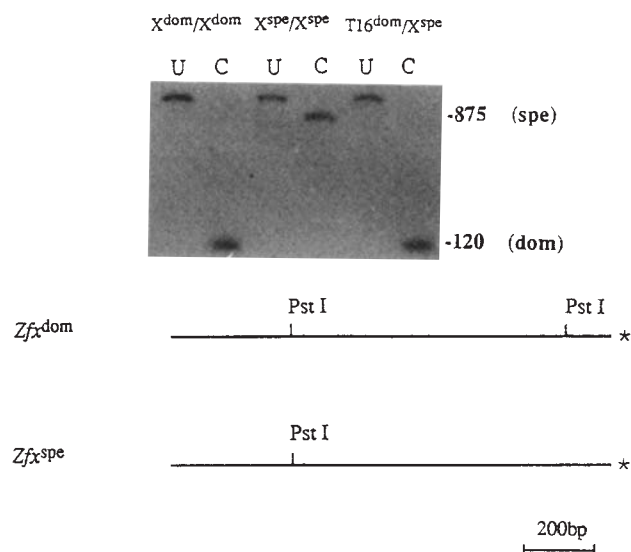
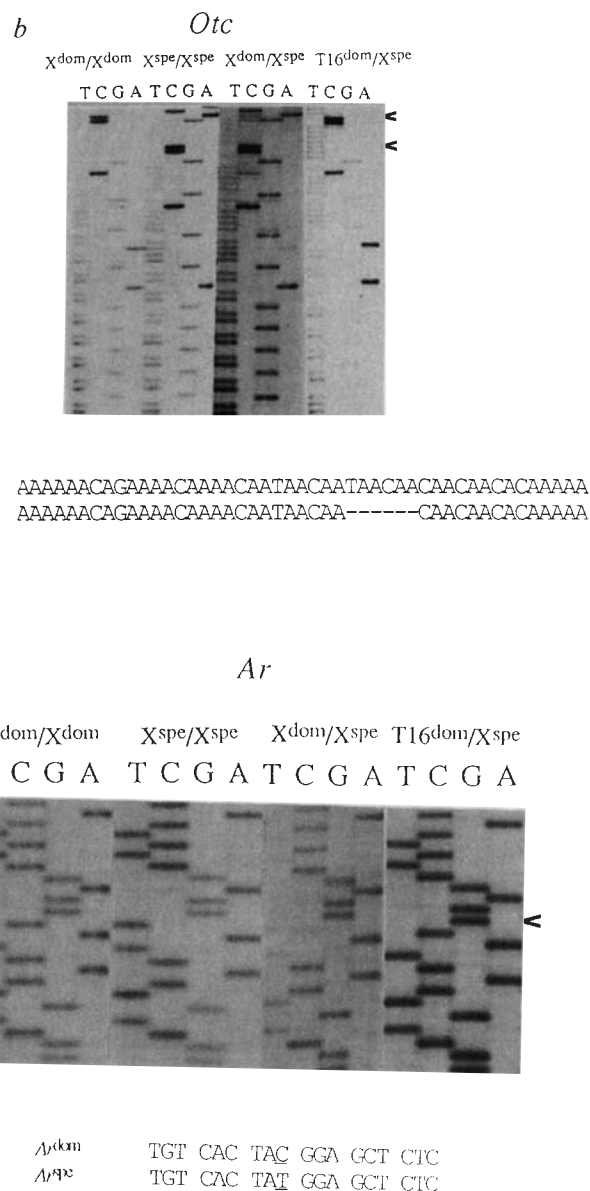


FIG. 2 Mouse *Zfx*, unlike human *ZFX*, undergoes X-inactivation. *Zfx* cDNA was amplified from the RNAs described in Fig. 1. One of the primers (indicated with an asterisk) was end-labelled with ³²P. A restriction enzyme site polymorphism was used to distinguish *Zfx*^{dom} from *Zfx*^{spe}. Digestion of *Zfx*^{dom} cDNA with *Pst*I gives a labelled fragment of 120 bp, whereas *Zfx*^{spe} gives one of 875 bp. Only the *Zfx*^{dom} allele is expressed in the T16^{dom}/ X^{spe} animal, indicating that *Zfx* is subject to X inactivation. U indicates PCR product and C indicates PCR product after *Pst*I digestion.

METHODS. Primers were TTCATCTGGACTGACTATGG and CAAGAGCATGTTGTTGCACC corresponding to nucleotides 663–682 and 1,858–1,839 of the mouse *Zfx* cDNA¹⁹. The 3' primer was end-labelled with [γ -³²P]ATP and T4 polynucleotide kinase before the PCR. The primers were designed not to amplify the *Zfa* gene which is derived by retroposition of one of the alternatively spliced *Zfx* RNAs lacking nucleotides 632–781 of the published *Zfx* sequence^{18,19}. The *Pst*I polymorphism in *Zfx* cDNA was inferred from the Southern blotting data of Mitchell *et al.*²⁵ Completion of the restriction enzyme digest was monitored by inclusion of plasmid DNA containing a *Pst*I site (data not shown) and then the gel autoradiographed.

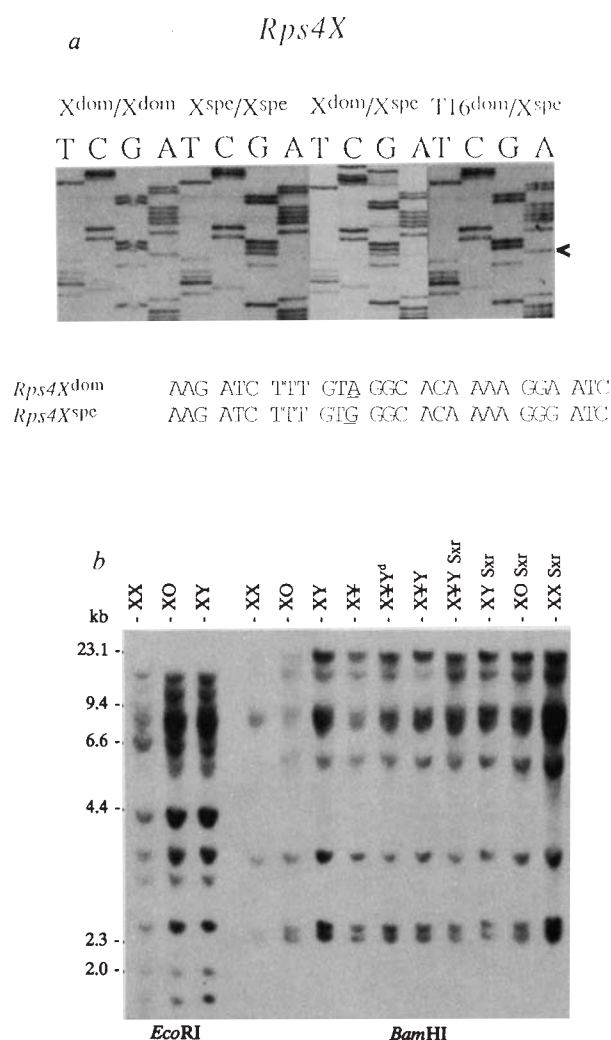


FIG. 3 Dosage compensation of *Rps4* is achieved by different mechanisms in mice and humans. *a*, *Rps4X* in mice is subject to X-chromosome inactivation. *Rps4X* cDNA was amplified from RNA from the various mice described in Fig. 1 and sequenced. A single-base variant (indicated by the arrowhead) was used to distinguish *Rps4X*^{dom} from *Rps4X*^{spe}. Although the X^{dom}/X^{spe} mouse expresses both *Rps4X*^{dom} and *Rps4X*^{spe}, the T16^{dom}/X^{spe} animal expresses only *Rps4X*^{dom}, indicating that *Rps4X* is subject to X-inactivation. *b*, Southern blot analysis showing absence of a Y-linked homologue of *Rps4X* in mice. All the bands seen can be attributed to *Rps4X* or to autosomal homologues. No Y-specific bands were revealed by digestion with two different restriction enzymes, or by looking at Y chromosome variants or dosage of Y chromosomal material.

METHODS. *a*, PCR techniques were as described in Fig. 1. One oligonucleotide (CGGCTCCAAAACACTGGATG) was synthesized based on the published rat *Rps4* sequence²⁶ (residues 40–59) and the other (TGGCAAAGCTGTTGCCA-TTG) from the sequence of a murine *Rps4X* cDNA clone (S. Swift and A.A., unpublished results). Sequencing was performed with a third primer (ATGG-ATGTCATCAGCATTG) complementary to the mouse sequence. *b*, Genomic DNA was prepared from mice with sex chromosome constitutions as indicated. Y^d represents a Y chromosome from the CD1 outbred line and is of *Mus m. domesticus* type, all other Y chromosomes and Sxr are of *Mus m. musculus* origin. Y, represents a Y chromosome that has been shown to carry a small deletion around *Sry* (see refs 27–29, and J. Gubbay and R. L.-B., unpublished observations). Mice with multiple Y chromosomes were obtained as described previously²⁸. About 10 µg DNA from each genotype was digested with *Eco*RI or *Bam*HI, electrophoresed on a 0.7% agarose gel and transferred to a Hybond N filter (Amersham). The filter was hybridized to a ³²P-labelled 600-bp murine *Rps4X* cDNA fragment (see above). After washing with 2 × SSC buffer at 65 °C, the filter was autoradiographed for 12 h at –70 °C.

by Southern blot analysis (Fig. 3), polymerase chain reaction analysis and complementary DNA cloning (data not shown). Thus it seems likely that XY, XX and XO mice all express equivalent amounts of *Rps4* after X-inactivation.

We have shown that both *Zfx* and *Rps4X* undergo X-inactivation in the mouse, but not in humans. But because of the absence of a Y-linked gene or gene product, the effect of X-inactivation of these genes in the mouse is to equalize the dosage of the genes between XX females and XY males. In humans where ubiquitously expressed Y-linked copies for both genes exist, dosage compensation between males and females is achieved

by escape from inactivation of the X-linked homologue. As a result of the difference in the way the dosage compensation is achieved, however, XO humans will suffer haploinsufficiency for *ZFX* and *RPS4X*, which may contribute to their poor viability. By contrast, XO mice will be equivalent to normal XX or XY mice in the level of these gene products. The evidence indicates that X inactivation is complete in the mouse, which suggests that the viability of XO mice and absence of a Turner phenotype is the result of the complete inactivation of critical X-linked genes and their absence or restricted expression from the Y chromosome. □

Received 13 February; accepted 23 April 1991.

- Turner, H. H. *Endocrinology* **23**, 566–574 (1938).
- Ferguson-Smith, M. A. *J. med. Genet.* **2**, 142–155 (1965).
- Burgoyne, P. *Nature* **342**, 860–862 (1989).
- Fisher, E. M. C. *et al. Cell* **63**, 1205–1218 (1990).
- Lyon, M. F., Searle, A. G., Ford, C. E. & Ohno, S. A. *Cytogenetics* **3**, 306–323 (1964).
- Bonhomme, F. & Guenet, J.-L. in *Genetic Variants and Strains of the Laboratory Mouse* 2nd edn (eds Lyon, M. & Searle, A. G.) 649–662. (Oxford University Press, 1989).
- McMahon, A. & Monk, M. *Genet. Res.* **41**, 69–83 (1982).
- Rastan, S. *J. Embryol. exp. Morph.* **78**, 1–22 (1983).
- Cattanach, B. M. & Johnston, P. *Hereditas* **94**, 5 (1981).
- Johnston, P. G. & Cattanach, B. M. *Genet. Res.* **37**, 151–160 (1981).
- Rastan, S. *Genet. Res.* **40**, 139–147 (1982).
- Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. & Page, D. C. *Cell* **57**, 1247–1258 (1989).
- Brown, C. J. & Willard, H. F. *Am. J. hum. Genet.* **45**, 592–598 (1989).
- Page, D. C. *et al. Cell* **51**, 1091–1104 (1987).
- Palmer, M. S. *et al. Nature* **342**, 830–836 (1989).

- Koopman, P., Ashworth, A. & Lovell-Badge, R. *Trends Genet.* **7**, 132–136 (1991).
- Koopman, P. *et al. Nature* **342**, 940–942 (1989).
- Ashworth, A., Skene, B., Swift, S. & Lovell-Badge, R. H. *EMBO J.* **9**, 1529–1534 (1990).
- Mardon, G. *et al. Molec. cell. Biol.* **10**, 681–688 (1990).
- DeMars, R., LeVan, S. L., Trend, B. L. & Russell, L. B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1693–1695 (1976).
- Tettenborn, U., Dofuku, R. & Ohno, S. *Nature* **234**, 37–40 (1971).
- Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).
- Vere, G., Gibbs, R. A., Scherer, S. E. & Caskey, C. T. *Science* **237**, 415–417 (1987).
- Chang, C., Kokontis, J. & Liao, S. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7211–7215 (1988).
- Mitchell, M. *et al. Genetics* **121**, 803–809 (1989).
- Devi, K. R. G., Chan, Y.-L. & Wool, I. G. *Biochim. biophys. Acta* **1008**, 258–262 (1989).
- Lovell-Badge, R. H. & Robertson, E. *Development* **109**, 635–646 (1990).
- Gubbay, J. *et al. Development* **109**, 647–653 (1990).
- Gubbay, J. *et al. Nature* **346**, 245–250 (1990).

ACKNOWLEDGEMENTS. We thank P. Burgoyne for discussions and access to unpublished results and H. King for rapid oligonucleotide synthesis. This work was supported by the Cancer Research Campaign and the MRC.

Pre-existent pattern in *Xenopus* animal pole cells revealed by induction with activin

S. Sokol & D. A. Melton

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

ACTIVIN, a peptide growth factor related to tumour growth factor- β , has been implicated in early inductive interactions in vertebrates¹⁻⁴ and can induce *Xenopus* blastula ectodermal explants to develop a rudimentary axial pattern with anteroposterior and dorsoventral polarity^{4,5}. Here we demonstrate that prospective dorsal and ventral regions of the ectoderm respond differently to the same concentration of activin. Thus, activin does not seem to endow ectodermal cells with polarity but rather reveals a pre-existent pattern. Our results suggest that patterning of mesoderm is determined not only by a localized inducer, but also by the differential competence of cells in the responding tissue.

The development of a body axis in *Xenopus* seems to arise by the action of a localized inducer or morphogen on a homogeneous population of responding cells (animal cap cells or presumptive ectoderm). For example, when animal cap cells are rotated 180°, the polarity of the resulting embryo depends only on the polarity of the inducing (vegetal) cells⁶. Similarly, transplantation of a Spemann organizer can induce ventral ectodermal and mesodermal cells to form dorsal structures⁷. Thus the animal cap may consist of naive cells because their fates can readily be changed in response to a localized inducer.

That animal caps can be induced to form embryoids with a rudimentary axial pattern following incubation in a uniform

solution of activin⁵ has led us to examine the origin of polarity in induced explants. To investigate whether all animal pole cells exposed to an inducer differentiate into similar structures, we isolated prospective dorsal or ventral ectoderm regions from *Xenopus* blastula (stage 8–8.5) and compared their response to induction by an activin A homologue, PIF, which is an inducing factor derived from the P388DI cell line (see ref. 4). The differentiation of prospective dorsal and ventral animal cap halves after induction takes different routes (Fig. 1; Table 1). Two days after exposure to PIF the majority of prospective dorsal cap halves develop a rounded head-like appearance; roughly half of these explants contain eyes. Histological sections reveal abundant neural tissue, muscle and notochord (Fig. 1; Table 1). In contrast, sections of ventral halves induced with PIF show the presence of epidermis, muscle and mesenchyme (Fig. 1) and only rarely notochord or eyes (Table 1). Even high ($\times 20$) concentrations of the inducer fail to induce dorsoanterior structures in animal cap ventral halves (data not shown). Control halves of animal caps incubated in the medium without inducer develop into atypical epidermis without any mesoderm (mesenchyme, muscle or notochord) or neural tissue (Table 1; refs 4, 5). These experiments show that prospective dorsal and ventral regions of blastula animal caps are predisposed to develop towards a particular pathway following induction.

When do animal cap cells acquire this predisposition or polarity? Caps were cut from early, mid- and late-blastula stages, dorsal and ventral halves were separated and incubated in PIF. Dorsal halves separated at early blastula develop eyes as often as those isolated from late blastula (Table 1). Ventral halves isolated from early or late blastula rarely form eyes. Thus, the competence of prospective dorsal animal blastomeres to form eyes is established by the early blastula stage (stage 6.5–7). This dorsoanterior specification occurs long before the activity of a Spemann organizer can be demonstrated (late blastula, stage 9)⁸.

FIG. 1 Differential response of prospective dorsal (D) and ventral (V) halves of animal caps to PIF, which corresponds to mouse activin⁴. Prospective ventral (a, c) or dorsal (b, d) regions of animal caps were cut at blastula (stage 7), treated with PIF and cultured for 2 days. Morphology (a, b) and histological sections (c, d) of the induced explants reveal the presence of eyes, e; cement gland, cg; neural tissue, n; and notochord, nc, in dorsal halves of animal caps. Ventral halves of animal caps formed principally mesenchyme and muscle, m. **METHODS.** Eggs were obtained from *Xenopus laevis* females and embryos were raised as described¹⁹ and staged according to Nieuwkoop and Faber²⁰. Blastula animal caps were cut from blastulas about 150 μ m above the equator. None of the uninduced (control) half-caps contained any neural or mesodermal tissue as judged by histology (not shown). Only 6% of the uninduced control half-caps showed slight elongation after culturing ($n=35$). Preparation of PIF-containing medium and animal cap induction assays were performed as described⁵. The prospective dorsoventral axis was determined by pigmentation in the early embryo²⁰. Prospective ventral blastomeres are more heavily pigmented than their dorsal counterparts. The accuracy of this determination was tested in each experiment by leaving a large control group of embryos to develop to see if the pigmentation differences correctly predicted the position of dorsal blastopore lip. Explants were fixed in Bouin's solution, embedded in Paraplast, sectioned at 7 μ m and stained with haematoxylin–eosin. Scale bars, a and b, 80 μ m; c and d, 15 μ m.

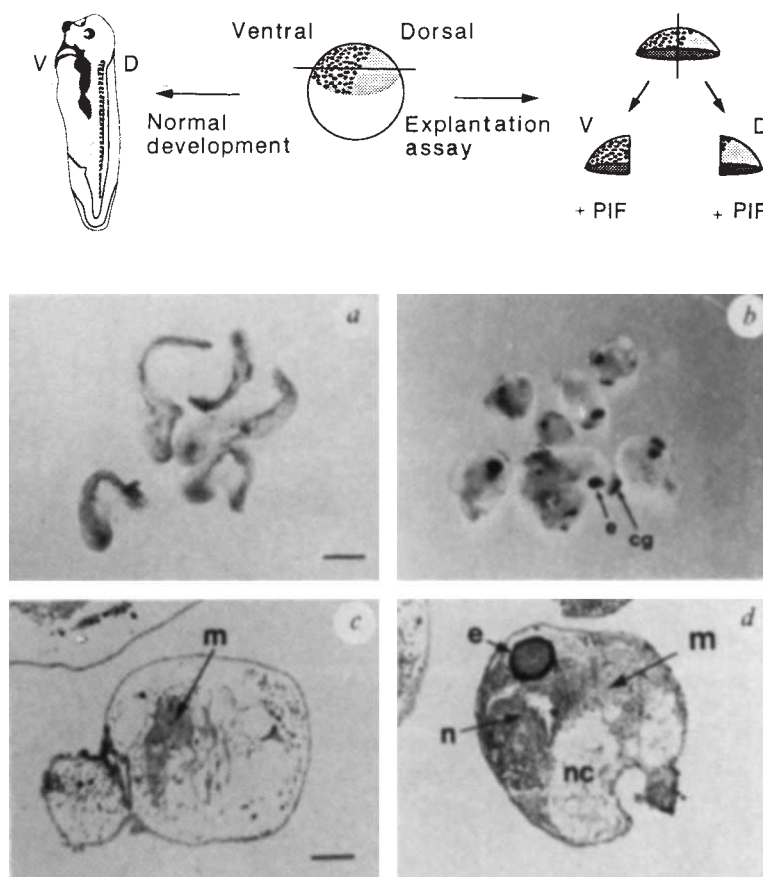


TABLE 1 Response of ectoderm regions to induction by PIF

Type of explant	6.5-7		Stage of explant 8-8.5				9-9.5	
	D	V	D	V	D	V	D	V
Inducer	+	+	-	-	+	+	+	+
Explants with eyes	41/86 (47%)	9/93 (9%)	0/22 (0%)	0/22 (0%)	53/102 (51%)	6/111 (5%)	36/89 (40%)	0/99 (0%)
Explants with notochord	ND	ND	0/12 (0%)	0/11 (0%)	15/17 (88%)	1/16 (6%)	ND	ND

Differential competence of prospective dorsal and ventral ectoderm to respond to PIF is established by early blastula stages. Prospective dorsal or ventral ectodermal regions from animal caps were explanted from three blastula stages of development and induced with PIF as described in Fig. 1. The differentiation of eyes is a marker for anterior neural development and differentiation of notochord is a marker for dorsal mesoderm. Results were scored after 2 days by counting the number of explants with induced eyes and by histological inspection for notochord. The data are presented as the ratio of explants with eyes or notochord over the total number of explants. Control animal caps incubated without inducer did not form mesodermal tissues (notochord, muscle or mesenchyme). ND, not determined; D, dorsal; V, ventral.

Lineage tracer (fluoresceinated dextran) was injected into blastomeres on the presumptive dorsal side and animal caps were cut from the injected embryos and treated with PIF. The differentiated explants contained lineage tracer in notochord, neural tissue, eye retina and muscle (Fig. 2). The lineage tracer only rarely appears in epidermis (8% of cases, $n=50$). In contrast, when lineage tracer is injected into an animal pole blastomere on the prospective ventral side, the tracer appears in epidermal, mesenchymal or muscle tissues (Fig. 2), but very

rarely in brain, other neural structures or notochord (6.2% of cases, $n=48$). All these experiments indicate that untreated animal caps possess a dorsoventral polarity. Cells in the prospective dorsal half of the cap differentiate into dorsoanterior structures when treated with PIF, and prospective ventral cells differentiate into ventrolateral structures. This polarity corresponds to the previously demonstrated dorsoventral polarity of the vegetal hemisphere^{6,9}.

Ultraviolet irradiation blocks the cytoplasmic rotation that occurs in the first cell cycle and this results in ventralized embryos that lack dorsal axial structures¹⁰. To determine whether PIF can induce formation of dorsal structures in ventralized embryos, fertilized eggs were exposed to ultraviolet irradiation and animal caps were cut at the blastula stage and incubated in PIF. The complete absence of eyes ($n=53$; Fig. 3) and the failure to form notochord (data not shown) shows that dorsoanterior induction is suppressed. The control group of animal caps cut from untreated embryos and treated with PIF developed eyes in 48% of the explants ($n=58$). These results suggest that an early ultraviolet-sensitive event is required for establishing animal cap competence to respond to PIF and develop dorsoanterior structures.

Our data show that PIF cannot induce formation of dorsoanterior structures in prospective ventral ectoderm of normal embryos nor in caps isolated from ultraviolet-irradiated embryos. As transplantation of dorsovegetal blastomeres from a normal 32-cell stage embryo induces a dorsal axis in ultraviolet-irradiated embryos¹¹, we propose that activin alone is not equivalent to the inducing signal(s) provided by dorsovegetal cells. This is further supported by the fact that an organizer graft can induce a dorsal axis in the ventral tissue of recipient embryos⁷.

In principle, embryonic polarity could be determined either by the action of a localized inducer on a homogeneous cell population or by differential competence of responding cells to

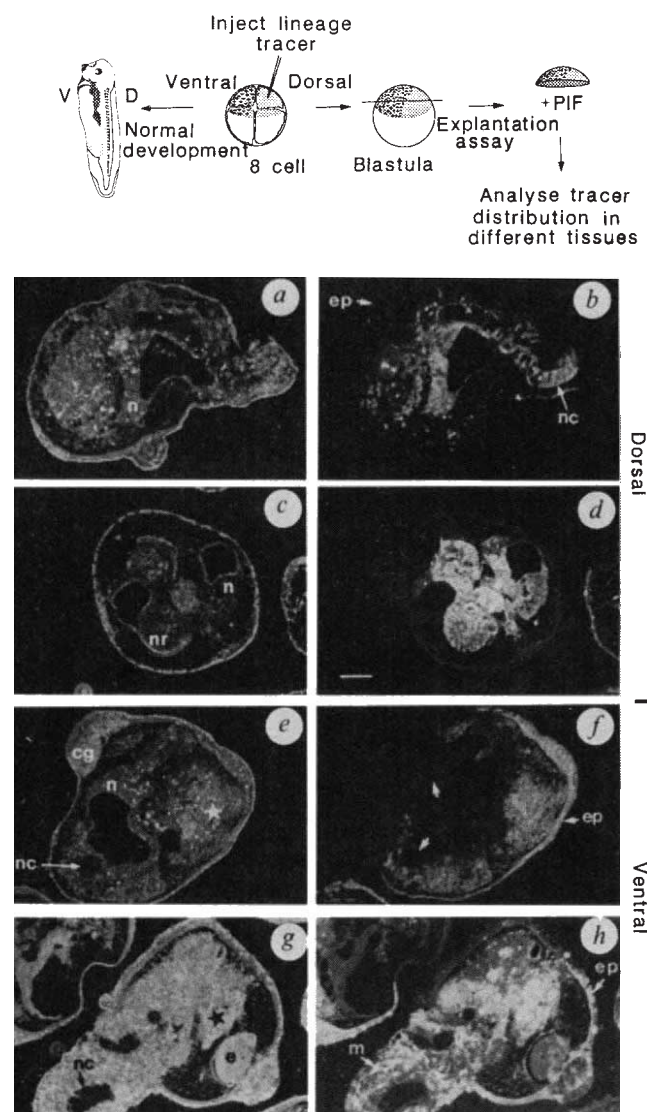


FIG. 2 Lineage tracing experiments reveal a predisposition of animal cap cells to form dorsoanterior or ventrolateral tissues following induction by PIF. Fertilized eggs developed to the 8-16 cell stage, at which time they were injected with fluoresceinated dextran into prospective dorsal (D; a-d) or ventral (V; e-h) animal blastomeres. Animal caps were cut from the injected embryos at the blastula stage (stage 7-7.5) and incubated with PIF. After 2 days, induced explants were sectioned and examined under ultraviolet light with epifluorescent optics. a, c, e, g, Darkfield view; b, d, f, h, fluorescence. After injection into a prospective dorsal blastomere, fluoresceinated dextran is later found in notochord, neural tissue and muscle, but not in epidermis (ep). Ventrally injected fluoresceinated dextran is found in epidermis and muscle, but not in neural tissue or notochord. Star denotes large cells with unclear tissue type. Scale bar, 75 μ m. METHODS. Embryos were injected with 5 nl solution of 10 mg ml⁻¹ fluoresceinated dextran (a gift from R. Gimlich). The prospective position of the dorsoventral axis was determined as described in the legend to Fig. 1. Abbreviations as in Fig. 1, except nr is neural retina.

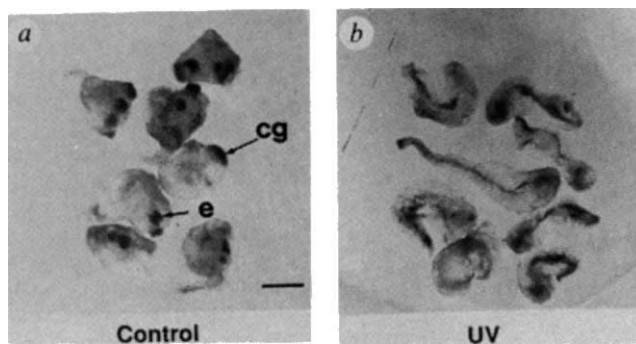


FIG. 3 PIF does not rescue dorsoanterior development in animal caps isolated from ultraviolet-irradiated embryos. Embryos were irradiated with ultraviolet light before first cleavage¹⁰. The effectiveness of ultraviolet irradiation (complete dorsal deficiency) was verified by examining the phenotypes of ultraviolet-irradiated embryos at later stages. Animal caps were cut at the blastula stage (stage 7–8) and treated with PIF. Groups of induced caps from irradiated (b) and untreated (a) embryos are shown. Note that PIF-induced caps, cut from the irradiated embryos, are indistinguishable from prospective ventral halves treated with PIF (compare with Fig. 1b). Histological sections reveal that animal caps derived from irradiated embryos and induced with PIF do not form notochord (not shown). Abbreviations and methods as in the legend to Fig. 1. Scale bar, 250 μ m.

a homogeneously distributed inducer. In the first case, differences in the local concentration of an inducing agent would establish the pattern. In support of this view, low concentrations of activin added to dispersed animal cap cells induce both muscle and notochord, whereas high concentrations give rise to cells with neural inducing capacity¹². Notably, co-induction of muscle and notochord occurs at the same concentration of activin¹². Although these results show that activin exhibits some properties of an instructive morphogen, they do not fully explain how differences between dorsal (notochord) and ventrolateral (muscle and mesenchyme, for example) mesoderm are established. Our results suggest that the dorsoventral character of mesoderm is not determined solely by the concentration of the inducer (activin), but also depends on the position of the responding cells in the embryo.

That presumptive dorsal and ventral ectoderm differ in the distribution of an epidermal marker¹³ has been interpreted to indicate a difference in competence to neural induction¹⁴. We find that dorsoventral differences in the presumptive ectoderm are important in mesoderm induction and that these differences are established very early in development, not later than the early blastula stage. These results are consistent with previous reports on the differential response of gastrula ectoderm to neural induction^{15,16}, a difference that may derive from the polarity described here.

We conclude that the axial polarity observed in PIF-induced embryoids derives from a pre-existent pattern of animal cap cells. This finding and other studies^{11,17,18} suggest that dorsoventral differences in the mesoderm may be determined not only by the action of a localized inducer, but also by the differential competence of cells in the responding tissue. □

Received 14 November 1990; accepted 17 April 1991.

- Asashima, M. et al. *Roux's Arch. devl Biol.* **198**, 330–335 (1990).
- Smith, J. C., Prke, B. M. J., Van Nimmen, K. & Huylebroeck, D. *Nature* **345**, 729–731 (1990).
- van den Eijnden-Van Raaij, A. J. M. et al. *Nature* **345**, 732–734 (1990).
- Thomsen, G. T. et al. *Cell* **63**, 485–493 (1990).
- Sokol, S., Wong, G. G. & Melton, D. A. *Science* **249**, 561–564 (1990).
- Nieuwkoop, P. D. *Roux's Arch. devl Biol.* **163**, 298–315 (1969).
- Spemann, H. *Embryonic Development and Induction* (Yale University Press, New Haven, 1938).
- Nakamura, O. in *A Milestone of a Half-century from Spemann* (eds Nakamura, O. & Toivonen, S.) 179–220 (Elsevier/North Holland Biomedical, Amsterdam, 1978).
- Dale, L. & Slack, J. *Development* **100**, 279–295 (1977).
- Scharf, S. & Gerhart, J. C. *Devl Biol.* **79**, 181–198 (1980).

- Gimlich, R. L. & Gerhart, J. C. *Devl Biol.* **104**, 117–130 (1984).
- Green, J. B. A. & Smith, J. C. *Nature* **347**, 391–394 (1990).
- London, C., Akers, R. & Phillips, C. *Devl Biol.* **129**, 380–389 (1988).
- Savage, R. & Phillips, C. R. *Devl Biol.* **133**, 157–168 (1989).
- Sharpe, C. R., Fritz, A., De Robertis, E. M. & Gurdon, J. B. *Cell* **50**, 749–758 (1987).
- Dixon, J. E. & Kintner, C. *Development* **106**, 749–757 (1989).
- Mitrani, E. et al. *Cell* **63**, 495–501 (1990).
- Ruiz i Altaba, A. & Jessel, T. *Genes Dev.* **5**, 175–187 (1991).
- Newport, J. & Kirschner, M. *Cell* **30**, 687–696 (1982).
- Nieuwkoop, P. & Faber, J. *Normal Table of Xenopus laevis (Daudin)*, (North-Holland, Amsterdam, 1967).

ACKNOWLEDGEMENTS. We thank P. Vize, M. Ku and P. Wilson for reading the manuscript and A. Ruiz i Altaba for advice. This research was supported by the NIH (D.M.).

Dynamin-like protein encoded by the *Drosophila shibire* gene associated with vesicular traffic

Alexander M. van der Bliek & Elliot M. Meyerowitz

Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125, USA

TEMPERATURE-sensitive paralysis is the most striking defect of adult *Drosophila* carrying the *shibire* mutation¹. This is believed to be due to a reversible block of endocytosis, which prevents membrane cycling and thus depletes synaptic vesicles^{2,3}. The *shibire* mutation also affects many tissues outside the nervous system^{4–7}. We have now mapped and characterized the *shibire* gene. A 275-kilobase yeast artificial chromosome was subcloned into cosmids, among which the gene was then located by analysing with restriction-fragment length polymorphisms. A 15-kilobase fragment of wild-type DNA rescues the mutant phenotype and the sequence of two mutant alleles show differences with wild type, demonstrating that we have isolated the *shibire* gene. The gene encodes a protein that is highly similar to rat dynamin^{8,9}, 69% of the amino-acid sequence is identical. Dynamin is a GTP-driven mechanochemical enzyme related to mammalian mx-proteins¹⁰ and to the yeast *vps1* gene product¹¹. Because the *shibire* gene product and dynamin have extensive similarity, we propose that they are cognate homologues. Dynamin causes microtubules to slide along each other *in vitro*¹² and in extracts it is associated with a distinct, but so far uncharacterized, membrane fraction¹³. In light of the *shibire* phenotype, we suggest that these proteins provide the motor for vesicular transport during endocytosis.

Shibire has been mapped by recombination to position 1-52.1 (refs 1 and 4), which corresponds to bands 13F-14A1 of the X chromosome. This is included in a yeast artificial chromosome (YACDY922) generated as part of the *Drosophila* genome project¹⁴. We have subcloned this 275-kilobase (kb) chromosome into cosmids and constructed a physical map, which includes genes encoding the *Drosophila* homologue of the *myb* protein, GAPDH, the G-protein β subunit, the clathrin heavy chain and the *scalloped* protein (Fig. 1a). The arrangement of cosmids was confirmed with *NotI* and *SfiI* digests of genomic DNA and of the YAC. To locate the *shibire* gene, we used 10 restriction-fragment length polymorphisms (RFLPs) that distinguish *scalloped* (strain BL911) from *shibire* (*shi*^{ts1}) DNA in the 275 kb spanned by the YAC. We isolated 23 flies, which showed recombination between the *shibire* and the *scalloped* genes, detected as double-wild-type male progeny of trans-heterozygote mothers. DNA for RFLP analysis was extracted from lines derived from each of the recombinant flies. RFLPs distal to *scalloped* were all of the type found in *shibire* parents, because the wild-type *scalloped* gene was contributed by the *shibire* chromosome (Fig. 1b). Conversely, RFLPs near the *shibire* gene should all be of the *scalloped* type. RFLPs that lie between the two genes could have the recombination breakpoint on either

side, and indeed, both RFLP alleles were found in the collection of recombinants. By plotting the frequency of each RFLP against the map position we located *shibire* (Fig. 1b). The RFLP data precluded the possibility that *shibire* was one of the previously cloned genes in 13F-14A1.

So far, three genes have been found by hybridizing RNA blots with a series of restriction fragments that encompass 40 kb targeted by the RFLP plot (Fig. 1c). There is no obvious rearrangement in the genomic DNA of the two mutant alleles (*shi*^{ts1} and *shi*^{ts2}) that we have examined, nor is gene expression affected. One large transcription unit, a 15-kb gene encoding a 3.5-kb messenger RNA, was the most likely candidate. This is consistent with the large target that *shibire* provides for ethylmethanesulphonate (EMS) mutagenesis¹. Furthermore, it is at the intercept of the RFLP plot (Fig. 1b). We transformed the germ line of mutant flies with restriction fragments that included a wild-type copy of the 15-kb gene and excluded the flanking genes (Fig. 1c). Two lines, derived from independent integration events, were established. Both rescued the *shibire* phenotype: instead of paralysis at 27 °C the adult transformants are paralysed at 30–35 °C (the temperature at which 50% of the flies were standing after 1 min exposure). This behaviour resembles that of *shi*^{ts1}/+heterozygous flies at 38 °C (ref. 15). The semi-

dominance of *shi*^{ts1} shows that there is a dosage effect, compounded in the rescue experiment by different expression levels commonly observed when a gene is placed in a new chromosomal location. Nonetheless, rescue is unequivocal and demonstrates that we have identified the *shibire* gene.

Fragments of the *shibire* gene were used as probes to isolate complementary DNA clones derived from adult fly RNA. We determined the nucleotide sequence of the longest clones, which indicated that they encode a protein of 836 amino acids (Fig. 2). The first methionine of this long open reading frame follows several stop codons and conforms to the consensus for translation initiation¹⁶. We have also found two classes of variant cDNAs. The first class encodes a protein with an internal deletion of either 2 or 6 amino acids at position 640 (Fig. 2). The second class has an alternative 3'-end, which results in the last two amino acids of the encoded protein being replaced by a 49-amino-acid extension. These variants were derived from alternatively spliced RNAs and are not cloning artefacts, because each variant was isolated several times independently.

A search of the GenBank data base revealed striking similarity between the *shibire* protein and rat dynamin⁸ (69% amino-acid sequence identity; Fig. 2), and to a lesser extent mammalian mx proteins¹⁰ and the yeast *vps1* gene product¹¹ (27% and 44%,

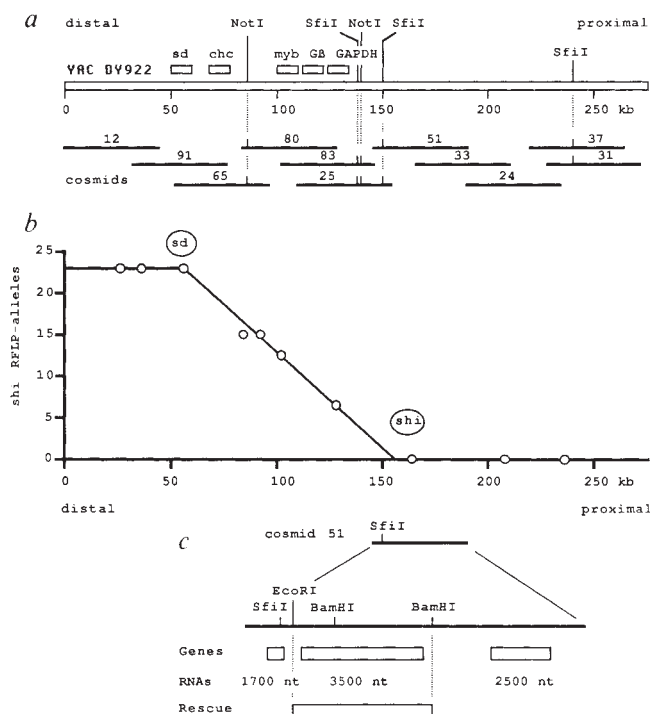


FIG. 1 a, Physical map of the genomic environment of *shibire*. The 275-kb YAC DY922 is depicted, showing the following genes: *scalloped* (*sd*), clathrin heavy chain (*chc*), *Drosophila myb* (*myb*), G protein β subunit (*Gβ*) and glyceraldehyde phosphate dehydrogenase (*GAPDH*). A representative series of cosmids isolated by subcloning the YAC is shown below. b, Plot of RFLPs in *scalloped-shibire* recombinant lines. The RFLPs (○) were used to detect the breakpoints between *shibire* (*shi*) and *scalloped* (*sd*). The number of recombinants with an RFLP allele derived from the *shibire* parental strain (coupled to a wild-type copy of the *scalloped* gene) is plotted against the position of the fragment. The intercept shows roughly the physical location of *shibire*. c, Enlargement of the RFLP-targeted area. The locations of the three transcription units are shown (genes), as are their transcript lengths (RNA) and the fragments used to rescue the *shibire* phenotype (rescue). Only the relevant restriction sites are shown.

METHODS. To subclone YAC DY922 in cosmids, yeast DNA was prepared in agarose and the chromosomes were size-fractionated with a CHEF-gel apparatus as specified by the manufacturer (Bio Rad). YAC-containing slabs were collected from five low-melting-point agarose gels, then melted at 65 °C, and cooled to 37 °C. Prewarmed 10 × *Sau3A* buffer (330 mM Tris-acetate, pH7.9, 660 mM potassium acetate, 100 mM magnesium acetate and 5 mM dithiothreitol) was gently mixed with the solution. Partial digests of 3 ml samples (about 2 µg DNA) were performed by diluting 0.01 or 0.03U *Sau3A* in 1/10 volume reaction buffer at 37 °C and gently mixing this with the DNA solution. The digest was terminated after 20 min by adding 1/10 volume 200 mM EDTA. The agarose was then digested by incubating twice for 2 h with 250U agarase at 37 °C, followed by a phenol extraction and ethanol precipitation. The DNA was dephosphorylated, ligated to the cosmid vector c2X75 (ref. 24), packaged and plated, yielding 50 recombinant cosmids derived from YACDY922. The arrangement of cosmids was verified with *NotI* and *SfiI* digests of YAC DNA and *Drosophila* genomic DNA embedded in agarose. Genes known to be in the area were located by hybridizing panels of the cosmids with the relevant probes. For the G-protein and *GAPDH* genes these were generated by polymerase chain reaction. RFLPs that allow the *scalloped* strain BL911 to be distinguished from *shi*^{ts1} were identified by hybridizing genomic digests with cosmids derived from the YAC. Recombinants between *scalloped* and *shibire* were generated by crossing the two mutant strains and searching for double wild-type F₂ males. Testing 14,000 males yielded 23 recombinants, which shows that the distance between *scalloped* and *shibire* is 0.3 cM, in agreement with previous measurements^{1,4}. Each RFLP was tested in all 23 recombinants and a plot was drawn schematically through the allele frequencies as shown. Transcription units were located in the area targeted by the RFLP-plot using genomic fragments to probe cDNA libraries and RNA blots. The largest transcription unit (a 15-kb gene) was recloned as a 5-kb *EcoRI*-*BamHI* and a 10-kb *BamHI* fragment in the vector *cas3*, which contains the *white* gene as marker flanked by P-element repeats²⁵. This construct, together with the transposase producing plasmid *phsπΔ2-3*, was injected into *w1118-shi*^{ts1} double-mutant embryos. Two lines representing independent integration events were established. Rescue of the *shibire* phenotype was tested by finding the temperature at which 50% of the flies are still standing after 1-min exposure.



FIG. 2 Alignment of shibire (shi) and dynamin (dyn) amino-acid sequences (single-letter amino-acid code). Identical amino acids are shown by vertical lines between the sequences. Above the shibire sequence are shown: the GTPase consensus (cons), sequences changed in *shi*^{ts1} (glycine to aspartate) and *shi*^{ts2} (glycine to serine) alleles and the position of introns (*). The splice variants are: a deletion of two ($\Delta 2$) or six amino acids ($\Delta 6$) at position 640 and the replacement of the C-terminal two amino acids by a 49 amino-acid extension.

METHODS. Probes derived from genomic fragments were used to screen a size-selected cDNA library from adult head RNA of a *Drosophila* Oregon-R

strain cloned in pEXLX²⁶. The longer cDNAs were sequenced in both strands using custom oligonucleotides as primers. At least two examples of each splice variant were sequenced. Exon boundaries were determined by sequencing the genomic fragment subcloned for microinjection (Fig. 1) with the same set of oligonucleotides. In each case, the sequences were found to conform with the rules for splice sites. The mutations were located by sequencing PCR products of cDNA synthesized with RNA from the *shi*^{ts1} and *shi*^{ts2} alleles. Two PCR clones per allele were sequenced and the mutated sequences were verified by repeating this process with PCR products of genomic DNA. The dynamin sequence was taken from ref. 8.

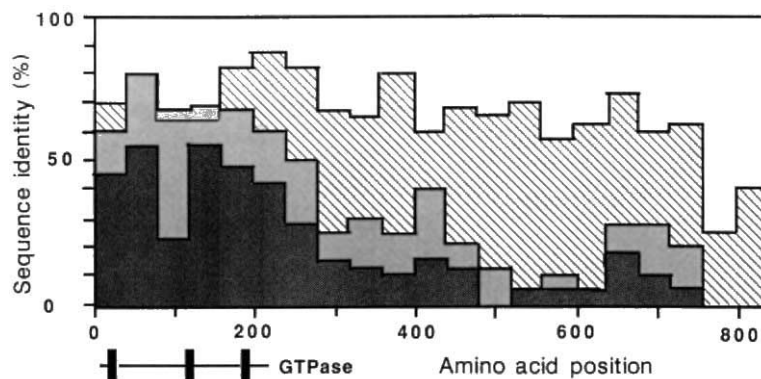
respectively). Each of these has a similar N-terminal GTPase domain. The similarity between the *shibire* gene product and dynamin remains surprisingly high throughout the rest of the protein, suggesting that they are cognate homologues (Fig. 3). Indeed, on isolating the human homologue of *shibire* by low-stringency hybridization, we found we had cloned the human dynamin gene (90% of the 500 sequenced base pairs were identical to rat dynamin; A.M.v.d.B., L. S. McEntee and E. M. M., unpublished data). The high degree of conservation between the *Drosophila shibire* gene product and mammalian dynamin suggests that they have the same basic function.

The mutations of two *shibire* alleles (*shi*^{ts1} and *shi*^{ts2}) were determined by sequencing cDNA and confirmed by sequencing

genomic DNA. The sequences provided additional evidence that we have identified the *shibire* gene. Both alleles have G-to-A transitions (common among EMS mutations), which change glycine residues to aspartate (*shi*^{ts1}) or serine (*shi*^{ts2}) (Fig. 2). Both glycines are in or near the GTPase domain and conserved in dynamin and the *vps1* protein. In mx proteins, which are the most distant relatives of this family, the corresponding glycines are also present, but they are shifted one position in the aligned sequences. Substitution of glycines near the GTPase consensus sequences suggests that this domain is disabled by a change in tertiary structure at the restrictive temperature.

The phenotypic data on *shibire* can now be brought to bear on the function of dynamin. The force-generating properties of

FIG. 3 Histogram of amino-acid sequence identity within the dynamin family. The percentage identity with the *shibire* protein was determined in blocks of 40 amino acids. The comparison between *shibire* and *dynamin* gene products is hatched, between *shibire* and *vps1* gene products shaded lightly and between *shibire* and *mx1* gene products shaded darkly. Each gap necessary for alignment was considered to be a sequence difference. From position 80 to 160 the percentage identity between *shibire* and *vps1* gene products was slightly higher than between *shibire* and dynamin proteins (superimposed as a thin line on the light shading).



dynamins have been demonstrated with microtubules sliding along each other *in vitro*¹². In extracts of rat brain, dynamin copurifies with a synaptosomal membrane fraction distinct from synaptic vesicles¹³. This is consistent with a role in endocytosis as suggested by the effect of *shibire in vivo*^{6,7,17,18}. For example, in temperature-sensitive mutants of *shibire* the number of intracellular vesicles at synapses decreases at the nonpermissive temperature, a process compounded by axon firing^{2,3}. Lack of membrane cycling provides an explanation for temperature-sensitive paralysis, as well as the wide variety of conditional defects observed outside the nervous system⁴⁻⁷. We speculate that dynamin motorizes intracellular traffic as do dynein and

kinesin^{9,19}, but that its action is restricted to endocytosis. Perhaps it attaches directly to internalized ligand-receptor complexes²⁰ or to their adaptors²¹. The *vps1* gene product may also motorize transport, primarily of vacuole-bound cargo¹¹, and the *mx*-proteins may specialize in evicting virus-laden vesicles²². Recently, a second function of the *vps1* product was found; the gene is identical to *spo15*, which is necessary for the microtubule-dependent process of separating spindle-pole bodies during meiosis²³. The question of molecules, other than tubulin, interact directly with dynamin, *vps1*- and *mx*-proteins can now be addressed by genetic means. □

Received 5 March; accepted 15 April 1991.

- Grigliatti, T. A., Hall, L., Rosenbluth, R. & Suzuki, D. T. *Molec. gen. Genet.* **120**, 107-114 (1973).
- Poodry, C. A. & Edgar, L. J. *Cell Biol.* **81**, 520-527 (1979).
- Koenig, J. H., Kosaka, T. & Ikeda, K. *J. Neurosci.* **9**, 1937-1942 (1989).
- Poodry, C. A., Hall, L. & Suzuki, D. T. *Dev. Biol.* **32**, 373-386 (1973).
- Raper Hummon, M. & Costello, W. J. *Wilhelm Roux Arch. dev. Biol.* **197**, 383-393 (1988).
- Kessel, I., Holst, B. D. & Roth, T. F. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4968-4972 (1989).
- Narita, K., Tsuruhara, T., Koenig, J. H. & Ikeda, K. *J. cell. Phys.* **141**, 383-391 (1989).
- Obar, R. A., Collins, C. A., Hammarback, J. A., Shpetner, H. S. & Vallee, R. B. *Nature* **347**, 256-261 (1990).
- Vallee, R. B. & Shpetner, H. S. *A. Rev. Biochem.* **59**, 909-932 (1990).
- Staeheli, P., Haller, O., Boll, W., Lindenman, J. & Wasserman, C. *Cell* **44**, 147-158 (1986).
- Rothman, J. H., Raymond, C. K., Gilbert, T., O'Hara, P. J. & Stevens, T. H. *Cell* **61**, 1063-1074 (1990).
- Shpetner, H. S. & Vallee, R. B. *Cell* **59**, 421-432 (1989).
- Scaife, R. & Margolis, R. L. *J. cell Biol.* **111**, 3023-3033 (1990).
- Garza, D., Ajioka, J. W., Burke, D. T. & Hartl, D. L. *Science* **246**, 641-646 (1989).
- Kim, Y.-T. & Wu, C.-F. *J. Neurogenet.* **7**, 1-14 (1990).
- Kozak, M. *Nucleic Acids Res.* **15**, 8125-8148 (1987).

- Masur, S. K., Kim, Y.-T. & Wu, C.-F. *J. Neurogenet.* **6**, 191-206 (1990).
- Poodry, C. A. *Dev. Biol.* **138**, 464-472 (1990).
- Vale, R. D. & Goldstein, L. S. B. *Cell* **60**, 883-885 (1990).
- Hopkins, C. R., Gibson, A., Shipman, M. & Miller, K. *Nature* **346**, 335-339 (1990).
- Pearse, B. M. F. *EMBO J.* **7**, 3331-3336 (1988).
- Arnheiter, H., Skuntz, S., Noteborn, M., Chang, S. & Meier, E. *Cell* **62**, 51-61 (1990).
- Yeh, E., Driscoll, R., Coltrera, M., Olins, A. & Bloom, K. *Nature* **349**, 713-715 (1991).
- Campbell, D. A. *Nucleic Acids Res.* **17**, 458 (1989).
- Pirrotta, V. in *Vectors: A Survey of Molecular Cloning Vectors and their Uses* (eds Rodriguez, R. L. & Denhardt, D. T.) 437-456 (Butterworths, Boston, 1988).
- Palazzo, M. J., et al. *Gene* **88**, 25-36 (1990).

ACKNOWLEDGEMENTS. We thank our laboratory colleagues and P. J. Johnson for advice and for reading the manuscript, E. Chang for technical assistance, B. A. Hamilton for the cDNA library, I. M. Duncan for the YAC, K. Vijay Raghavan, G. Payne and A. L. Katzen for probes, the *Drosophila* stock center (Bloomington, Indiana) for fly strains, and C. Bazinet for making us aware of the proximity of the clathrin gene. A.M.v.d.B. was supported by EMBO and the Human Frontier Science Program. This work was supported by an NIH Program Project (to E.M.M.). The *shibire* nucleotide sequence has been deposited with EMBL under accession numbers X59435 and X59436.

Membrane protein association by potential intramembrane charge pairs

Pierre Cosson, Scott P. Lankford, Juan S. Bonifacio & Richard D. Klausner

Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

THE transmembrane domain of the α chain of the T-cell receptor is responsible both for its assembly with the CD3 δ chain¹ and for rapid degradation of the unassembled chain within the endoplasmic reticulum^{2,3}. The determinant for both assembly and degradation is located in a segment of eight residues containing two basic amino acids (Fig. 1). We show here that placement of a single basic residue in the transmembrane domain of the Tac antigen can induce interaction with the CD3 chain, through its transmembrane acidic residue. This interaction is most favoured when the interacting residues are located at the same level in the membrane. The ability to induce protein-protein interaction by placing charge pairs within transmembrane domains suggests an approach to producing artificial dimers.

To investigate the role of the basic residues in the transmembrane domain of the T-cell receptor α chain (TCR α) in its interaction with the CD3 δ chain (CD3- δ), we used a chimaeric protein of the Tac antigen⁴ whose transmembrane domain was replaced by the corresponding domain of TCR- α (Tac α , Fig. 1). This chimaera and specific variants of it were cotransfected with CD3- δ into COS cells, and the level of assembly was determined by coprecipitation with either anti-Tac or anti- δ antibodies after metabolic labelling. Although full-length Tac did not interact with CD3- δ (Fig. 2a), the chimaeric Tac α demonstrated assembly, as shown by coprecipitation of CD3- δ with anti-Tac antibodies (Fig. 2b). Replacing either of the two basic amino acids with leucine did not markedly affect the level of interaction (Fig. 1c, d), but replacement of both basic residues

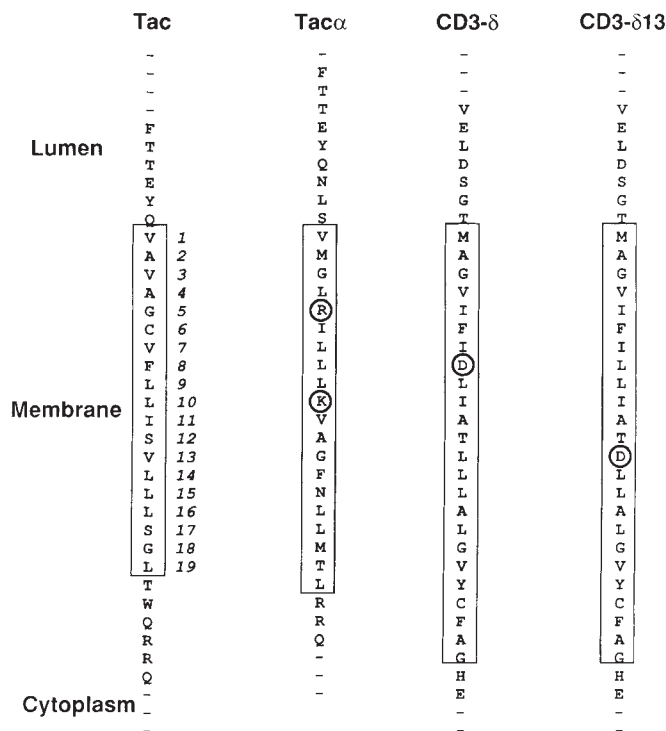


FIG. 1 Transmembrane domains of Tac, Tac α , and CD3- δ mutants. The sequences of the transmembrane domain and surrounding region of Tac (α chain of the human interleukin-2 receptor⁴), Tac α (a fusion protein composed of the Tac luminal domain, TCR- α transmembrane domain and Tac cytoplasmic domain), CD3- δ (ref. 11) and CD3- δ 13 are shown (single-letter amino-acid code). Boxed sequences correspond to the predicted transmembrane domains¹². The amino-acid residues of the transmembrane domain are arbitrarily numbered from the luminal to the cytoplasmic face. Potentially charged amino-acid residues are circled.

METHODS. Tac α (Tac-TCR- α 2) and Tac α mutants were described previously³. The Tac mutants with one arginine residue in their transmembrane domains were obtained by M13 mutagenesis, using the mutagenesis kit from BioRad. δ 13-Gal was obtained by mutagenesis using the polymerase chain reaction¹³.

reduced binding to a level only slightly higher than background (Fig. 1e).

To assay the assembly of multiple constructs, we made a fusion protein of CD3- δ with β -galactosidase (δ -Gal), where β -galactosidase was introduced in-frame at the C terminus of CD3- δ . After cotransfection, about 15% of the galactosidase activity was coprecipitated with Tac α . The results for the δ -Gal construct were qualitatively identical to those obtained after metabolic labelling of the CD3- δ chain (Fig. 3); either of the two basic amino-acid residues in the transmembrane of Tac α was sufficient to induce assembly with δ -Gal, but replacement of both by leucine decreased assembly by 70% (after subtraction of the background). We also tested the transmembrane domain of the TCR clonotypic δ chain (Ti δ) which contains the same basic amino acids as TCR- α , at the same predicted positions, but shares few other amino acids in this region⁵. Tac-Ti δ bound to δ -Gal with the same efficiency as Tac α (Fig. 3).

The presence of a transmembrane acidic residue in CD3- δ suggested that it was the interaction between that residue and a basic residue in the transmembrane domain of Tac α that was responsible for assembly. Indeed, no binding was observed

between Tac α and CD3- δ when the aspartic acid in CD3- δ transmembrane was mutated to a leucine residue (N. Manolios *et al.*, manuscript in preparation). Replacing both basic amino-acid residues in the Tac α transmembrane domain with acidic residues reduced binding to background levels (Fig. 3). To test the hypothesis that a single potential charge pair could mediate protein assembly, we constructed mutants of Tac in which single residues within the transmembrane domain were replaced by arginine. We found that a single arginine in the transmembrane domain of Tac can cause interaction with δ -Gal (Fig. 4, filled squares). The assembly was strictly dependent on the position of the arginine in the transmembrane domain; optimal binding occurred when the arginine was located at position 8 (Fig. 1), which is at the same predicted level as the aspartic acid in the CD3- δ transmembrane domain. The efficiency of binding decreased markedly as the arginine was moved away from this position.

We next constructed a mutant of CD3- δ , where the aspartic acid at position 8 was replaced by a leucine, and the leucine at position 13 was replaced by an aspartic acid (δ 13-Gal). The maximum association was the same as that observed for the

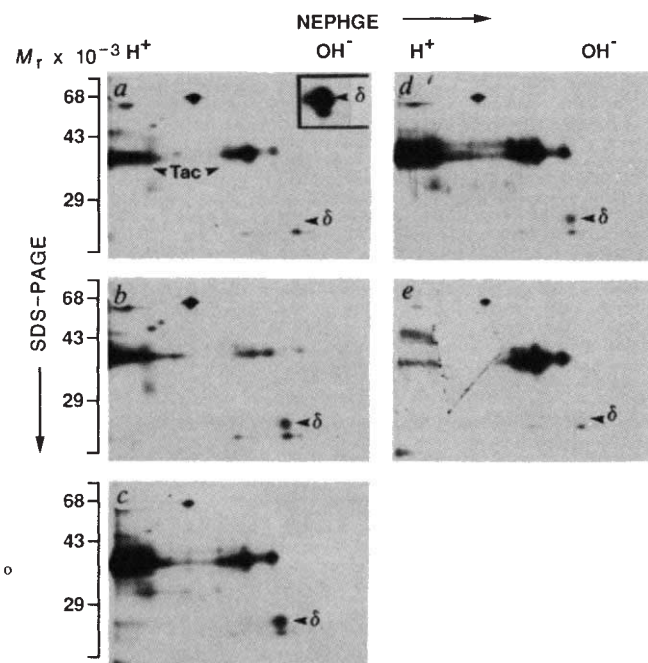


FIG. 2 Assembly of Tac α mutants with CD3- δ . COS-1 cells were cotransfected with expression plasmids encoding: a, Tac and CD3- δ ; b, Tac α and CD3- δ ; c, Tac α (Arg 5 $\rightarrow$ Leu) and CD3- δ ; d, Tac α (Lys 10 $\rightarrow$ Leu) and CD3- δ ; and e, Tac α (Arg 5 $\rightarrow$ Leu, Lys 10 $\rightarrow$ Leu) and CD3- δ . As previously described¹, 48 h after transfection, cells were metabolically labelled for 1 h at 37 °C. Cells were removed from the plates, solubilized with lysis buffer containing 0.5% (w/v) Triton X-100, centrifuged for 15 min at high speed, and supernatants were added to antibody beads. The immunoprecipitated proteins were resolved by two-dimensional non-equilibrium pH gel electrophoresis (NEPHGE)/SDS-PAGE. a-e, Immunoprecipitates with antibody to Tac. Inset in a shows the amount of total CD3- δ expressed in the same cells as analysed by immunoprecipitation with antibodies to CD3- δ . Identical amounts of CD3- δ were observed in transfections a-e. The positions of relative molecular mass (M_r) markers (expressed as $10^{-3} \times$ mass) are indicated on the left. The positions of Tac mutants (Tac) and CD3- δ (δ) are indicated. Experiments where the labelling time was reduced to 15 min gave similar results.

METHODS. Electrophoretic techniques have been described previously¹⁴. Antibodies to CD3- δ , directed against a C-terminal peptide of the murine δ chain¹⁵, were affinity-purified and covalently bound to Sepharose-4B beads activated by cyanogen bromide. The 7G7 antibody was directed against an epitope localized to the extracellular domain of the human Tac antigen¹⁶.

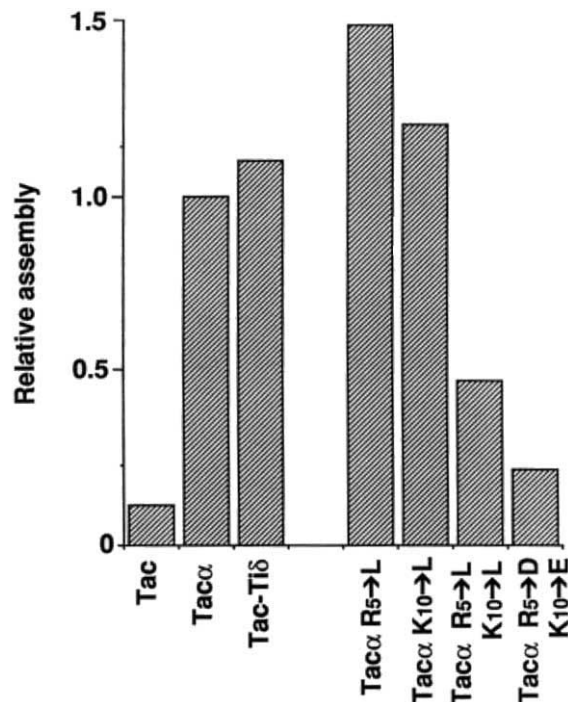
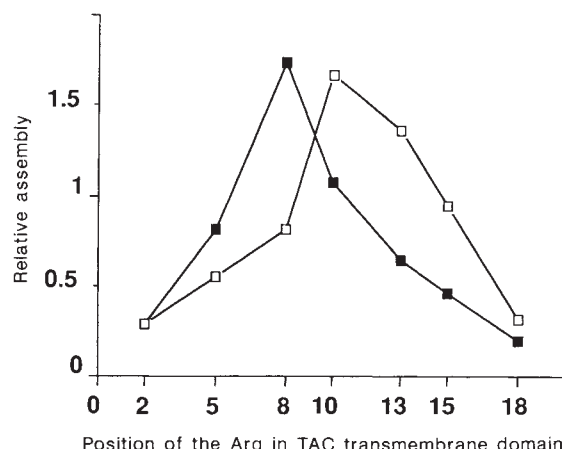


FIG. 3 Assembly of Tac α mutants with δ -Gal. COS-1 cells were cotransfected with plasmids encoding δ -Gal and the indicated Tac mutants. As in Fig. 1, the cells were lysed and the lysate precipitated with either anti-Tac or anti- δ antibodies. The galactosidase activity found in the anti-Tac immunoprecipitates was compared with the total galactosidase activity found in the anti- δ immunoprecipitate. In a typical cotransfection experiment with Tac α , 15% of the galactosidase activity is immunoprecipitated with anti-Tac antibodies. Relative assembly represents the ratio of the levels of assembly observed for each mutant and that observed with Tac α . K, Lys; L, Leu; D, Asp; R, Arg; E, Glu.

METHODS. δ -Gal is a fusion protein obtained by introducing at the end of the cytoplasmic domain of CD3- δ a linker of five alanine residues and the gene coding for the β -galactosidase enzyme. This fusion protein can be immunoprecipitated with an anti- δ antibody. To assay the galactosidase activity we used the substrate chlorophenol red- β -D-galactopyranoside (Boehringer Mannheim), and quantitated the product of the reaction by absorbance at 574 nm.

FIG. 4 Interaction between Tac mutants and δ -Gal: effect of the relative position of the two charged amino-acid residues. Tac mutants with a single arginine residue introduced at different levels in their transmembrane domains were cotransfected with either δ -Gal (■) or δ 13-Gal (□). As described in Fig. 3, the cells were lysed and the galactosidase activity was determined in an anti-Tac immunoprecipitate, as well as in the whole-cell lysate. The percentage of assembled δ - β Gal was calculated, and normalized to the levels obtained with Tac α . All experiments were repeated at least twice with different plasmid preparations and gave identical results to those presented here.



previous constructs, but the position for optimum binding was now between positions 10 and 13 (Fig. 4, open squares). These results suggest that the relative position of the interacting residues in the membrane determines the efficacy of assembly.

Whether assembly occurs by the actual formation of an ion pair or through hydrogen bonding remains unknown. The placement of a potentially charged residue in the lipid bilayer is an energetically unfavourable event^{6,7}. There is good theoretical support, however, for the energetic advantage of burying a charge pair in a hydrophobic environment over the transfer of two isolated charges^{7,8}. The placement of these residues within the hydrophobic region of the transmembrane domain probably allows a single potential charge pair to produce stable protein assemblies, and the energetic advantage of pairing should diminish as the residues are moved closer to the membrane interface. Indeed, when an arginine was placed in Tac at position 18, it did not mediate assembly with δ -Gal, nor with a δ 18-Gal (data not shown). The physical details of these interactions remain to be determined, including questions about bound ions or water, and their effects on the local structure of the membrane and whether other properties of the TM domains influence the assembly. Potentially charged residues in the membrane are

probably involved in the assembly of other oligomeric complexes, such as the B-cell antigen receptor⁹ or some Fc receptors¹⁰. It might also play a part in the folding of proteins with multiple membrane-spanning domains. □

Received 17 January; accepted 18 March 1991.

- Manolios, N., Bonifacio, J. S. & Klausner, R. D. *Science* **249**, 274–277 (1990).
- Bonifacio, J. S., Suzuki, C. K. & Klausner, R. D. *Science* **247**, 79–82 (1990).
- Bonifacio, J. S., Cosson, P. & Klausner, R. D. *Cell* **63**, 503–513 (1990).
- Leonard, W. J. *et al. Nature* **311**, 626–631 (1984).
- Raulet, D. H. *A. Rev. Immun.* **7**, 175–207 (1989).
- Engelman, D. M., Steitz, T. A. & Goldman, A. *Rev. Biophys. biophys. Chem.* **15**, 321–353 (1986).
- Parsegian, A. *Nature* **221**, 844–846 (1969).
- Honig, B. G. & Hubbell, W. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5412–5416 (1984).
- Wienands, J., Hombach, J., Radbruch, A., Riesterer, C. & Reth, M. *EMBO J.* **9**, 449–455 (1990).
- Kinet, J.-P. *Cell* **57**, 351–354 (1989).
- van den Eisen, P., Shepley, B. A., Cho, M. & Terhorst, C. *Nature* **314**, 542–544 (1985).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Higuchi, R., Krummel, B. & Saiki, R. K. *Nucleic Acid Res.* **16**, 7351–7367 (1988).
- Samelson, L. E., Harford, J. B. & Klausner, R. D. *Cell* **43**, 223–231 (1985).
- Samelson, L. E., Weissman, A. M., Robey, F. A., Berkowen, I. & Klausner, R. D. *J. Immun.* **137**, 3254–3258 (1986).
- Rubin, L. A., Kurman, C. C., Biddison, W. E., Goldman, N. D. & Nelson, D. L. *Hybridoma* **4**, 91–102 (1985).

ACKNOWLEDGEMENTS. P.C. is supported by the European Molecular Biology Organization. S.L. is an HMI-NIH Research Scholar.

Crystals of haemoglobin with the T quaternary structure bind oxygen noncooperatively with no Bohr effect

Andrea Mozzarelli, Claudio Rivetti, Gian Luigi Rossi, Eric R. Henry* & William A. Eaton*

Institute of Biochemical Sciences, University of Parma, 43100 Parma, Italy
*Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

THE relationship between the structure and function of haemoglobin has mainly been studied by comparing its X-ray crystal structures with its function in solutions^{1–11}. To make a direct comparison we have studied the functional properties of haemoglobin in single crystals, an approach that has been an important part of the investigation of several enzyme mechanisms^{12,13}. Here we report on the oxygen binding by single crystals of human haemoglobin grown in solutions of polyethylene glycol. Unlike haemoglobin crystals formed in concentrated salt solution, which crack and become disordered on oxygenation^{14–16}, crystals grown in polyethylene glycol remain intact. X-ray studies have shown

that the T (deoxy) quaternary structure of haemoglobin in this crystal at pH 7.0 is maintained at atmospheric oxygen pressure, and that the salt-bridges are not broken^{17–19}. We find striking differences between oxygen binding by haemoglobin in this crystal and by haemoglobin in solution. Not only is oxygenation of the crystal noncooperative, but the oxygen affinity is independent of pH in the range 6.0–8.5, and is much lower than that of the T state in solution. The lack of cooperativity without a change in quaternary structure is predicted by the two-state allosteric model of Monod, Wyman and Changeux^{1,5}. The absence of a Bohr effect without breakage of salt-bridges is predicted by Perutz's stereochemical mechanism^{2,4,9–11}. In contrast to the X-ray result that oxygen binds only to the α haems, our measurements show that the α haems have only a slightly higher affinity than the β haems.

Binding curves were obtained from polarized absorption spectra as a function of oxygen pressure on single crystals using a microspectrophotometer. X-ray diffraction of these crystals showed that they belong to the same space group (P2₁2₁2) with the same unit cell dimensions as those reported in the previous X-ray study¹⁹. The spectra were measured with linearly polarized light incident on the (010) face of the crystal and polarized parallel to either the *a* or *c* crystal axis. At each pressure the fractional saturation of the ferrous haems with oxygen was determined by fitting the observed spectra with a linear combination of oxyhaemoglobin (oxyHb), deoxyhaemoglobin (deoxyHb) and methaemoglobin (metHb) reference spectra. If

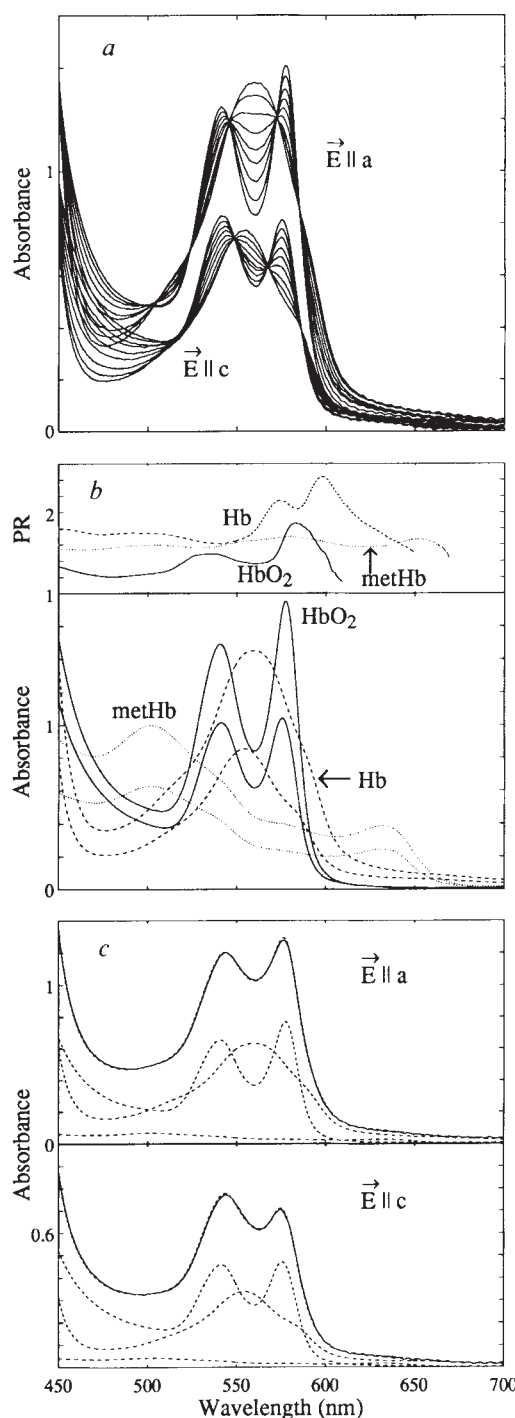
the concentration of polyethylene glycol in the external medium was not sufficiently high, on oxygenation crystals exhibited cracks, a decrease in polarization ratio and a time-dependent increase in oxygen affinity on a scale of minutes, signalling the conversion of the quaternary structure from T to R. All experiments reported below were done on crystals that remained intact with a time-independent affinity at all oxygen pressures.

Complete reversibility of the binding curve was established in an experiment at 15 °C in which the oxygen pressure was increased from 0 to 760 torr and back to 0 torr (Fig. 1a). The observation of nearly perfect isosbestic points in these crystal spectra shows that there is only one orientation for the oxyhaems and one orientation for the deoxyhaems, indicating that the T quaternary structure is maintained over the complete range of saturation. The small deviation from perfect isosbestic points

results from the change in the fraction of methaems and small baseline offsets during the course of the measurements which required 10 h to complete. Figure 1b shows the three reference spectra that were used to fit the observed spectra at each oxygen pressure, with a representative fit shown in Fig. 1c.

A Hill plot of these results shows that the association and dissociation curves are identical (Fig. 2). The Hill plot is linear up to the highest saturation of 83% with a slope of 1.00 ± 0.01 . A Hill n close to unity was found for all experiments in which the crystal remained intact (Fig. 3b). This result shows unambiguously that there is no cooperativity in oxygen binding by the T quaternary structure in the crystal. Lack of cooperativity in binding to the T quaternary structure is an essential feature of the two-state allosteric model of Monod, Wyman and Changeux^{1,5}.

FIG. 1 Single crystal polarized absorption spectra. Spectra were measured with a Zeiss MPM03 microspectrophotometer on single crystals covered with a gas-permeable membrane³⁵ and mounted in a flow cell. Oxygen pressures were measured with a Clark electrode. Additional details of the method will be presented elsewhere. A description of the measurement and analysis of polarized single crystal spectra of haemoglobin is presented in ref. 30. *a*, Single crystal spectra of human haemoglobin measured with linearly polarized light at 15 °C and pH 7.2 as a function of increasing oxygen pressure. The crystals were suspended in polyethylene glycol of relative molecular mass 8,000 (62% w/v, which is 125 g dissolved in 100 ml 10 mM potassium phosphate buffer). The upper set of spectra is for light polarized parallel to the *a* crystal axis, whereas the lower set of spectra is for light polarized parallel to the *c* crystal axis. *b*, Reference spectra of oxyHb, deoxyHb, and metHb used to fit the observed spectra. The polarization ratios (PR) at each wavelength, defined as the *a*-axis optical density divided by the *c*-axis optical density, are plotted above the spectra. The spectrum of metHb was measured after complete oxidation of a crystal with potassium ferricyanide. The spectrum of deoxyHb was obtained from the same crystal after washing and reduction with sodium dithionite. To obtain the oxyHb spectrum with no contribution from deoxyHb or metHb, the data at the highest pressures were extrapolated to infinite oxygen pressure in a separate experiment at 8 °C where the oxygen affinity is higher. The extrapolation used a straight-line fit to plots of the reciprocal of the Δ optical density (optical density at oxygen pressure p minus optical density at $p=0$) at each of the 251 wavelengths from 450 nm to 700 nm versus the reciprocal of the oxygen pressure. The contribution from metHb to the extrapolated spectrum was removed by subtracting increasing amounts of the metHb reference spectrum until the band at 630 nm disappeared. *c*, The observed spectrum at 157 torr and the fits with the reference spectra. The continuous curve is the observed spectrum, the short dashed curve is the sum of the component reference spectra (the fit is so good that the two spectra are almost indistinguishable), and the dotted curves are the component reference spectra. The metHb content at the beginning of this experiment was 2% and increased to 8% by the time that the system was equilibrated at 760 torr 10 h later. Top graph, light polarized parallel to *a* axis; bottom graph, light polarized parallel to *c* axis.



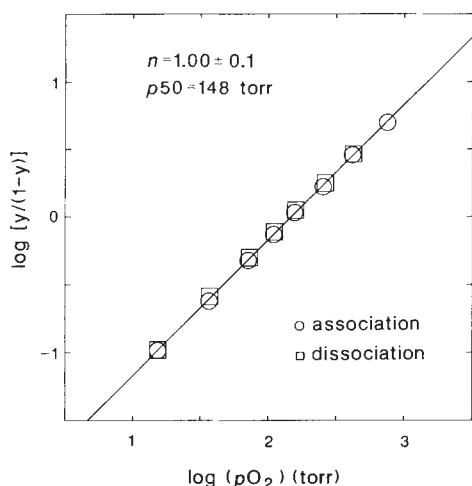


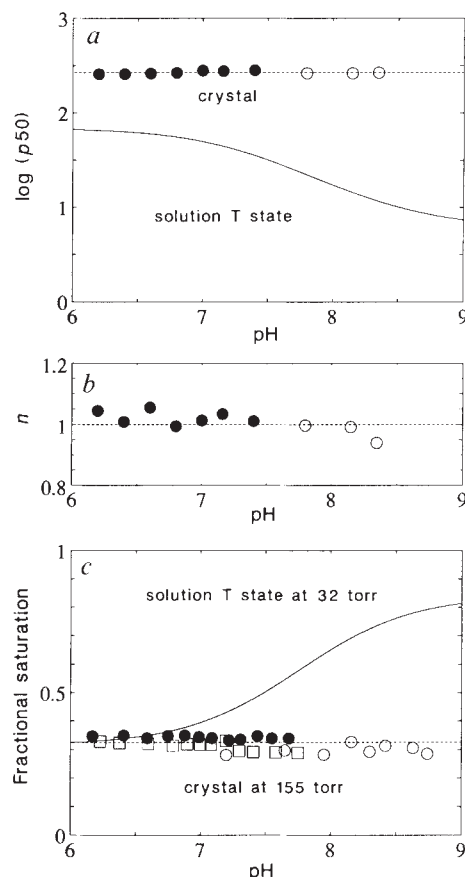
FIG. 2 Hill plot showing complete reversibility and non-cooperativity of the crystal binding curves. The fractional saturation with oxygen (y) is defined as the fractional saturation of the reduced (ferrous) haems (the fraction of oxyhaems divided by the sum of the fractions of oxyhaems plus deoxyhaems). $\circ$, Results from increasing the oxygen pressure from 0 to 760 torr calculated from the data in Fig. 1; $\square$, results from decreasing the oxygen pressure from 760 to 0 torr. The Hill n obtained by fitting all of the data is 1.00 ± 0.01 , with a $p50$ of 148 torr. The saturation at each oxygen pressure is the average of the saturations from the a -axis and c -axis data. The n values and $p50$ s calculated from the a and c data separately were 0.99 ± 0.01 , 155 ± 1 torr and 1.01 ± 0.01 , 141 ± 1 torr, respectively. The metHb content at the beginning was 2%, and at the end of the experiment, after increasing the oxygen pressure from 0 to 760 torr (Fig. 1) and back to 0 torr 20 h later, was 14%.

Both the oxygen affinity and the pH dependence of the oxygen affinity in the crystal are quite different from those found for solution. The oxygen pressure at 50% saturation ($p50$) for the crystal at 25 °C in 50 mM bis-Tris, 0.1 M NaCl, pH 7.4, is 270 torr, compared with 35–50 torr for oxygen binding to the T state in solution under very similar conditions^{20–23}. The more informative result is that the $p50$ for the crystal is independent of pH between 6.0 and 8.5 (Fig. 3a). This was confirmed in separate experiments which showed that the saturation of crystals in air is independent of both pH and chloride ion concentration (Fig. 3c). In contrast, in solution both the T-state affinity, assumed to correspond to the affinity for the first ligand (Fig. 3a, c) and the overall affinity, as measured by the $p50$,

increase with increasing pH and decreasing chloride ion concentration^{20,22,24,25}.

Assuming that the salt-bridges remain intact over the pH range of our experiments, the lack of a Bohr effect in this crystal is predicted by Perutz's stereochemical mechanism^{2,4,9–11,26}. In Perutz's mechanism, ligand binding to a subunit in the T quaternary structure, or changing the quaternary structure from T to R, produces a conformational change which breaks salt-bridges and releases Bohr protons. Prevention of both the tertiary and quaternary conformational changes by the lattice forces thus eliminates the Bohr effect. Presumably, prevention of the tertiary conformational change also lowers the affinity of the T quaternary structure compared with that in solution by not allowing

FIG. 3 Comparison of the pH dependence of T-state oxygen affinity in solution and in the crystal at 25 °C. **a**, Plot of $\log p50$ against pH. The points are the measured $p50$ s of the crystal from the spectra polarized parallel to the c crystal axis. The continuous curve is the data for the binding constant of the first oxygen molecule to haemoglobin, $\log(4/K_{41})$ (the statistical factor of 4 is missing from the label to Fig. 1a in ref. 20), where $K_{41}/4$ is assumed to be equal to the T state binding constant, K_T . Oxidation at 25 °C is relatively rapid, and the fraction of methaems at the end of the titrations varied from 10% at the highest pH values to 40% at the lowest pH values. A separate series of experiments showed, however, that oxidation has no effect on the affinity. In the X-ray study it was estimated that up to 30% of the β haems were oxidized⁹. **b** Plot of n against pH. These are the slopes of the Hill plots. **c**, Saturation versus pH at constant oxygen pressure at 25 °C. The points are the saturations at 155 torr oxygen, and the continuous curve is the saturation calculated at 31 torr oxygen from the values of $K_{41}/4$ of Lee *et al.*²⁰ using the Hill equation with $n=1$. The saturations are the average of those determined from the a and c crystal spectra. $\bullet$, 50 mM bis-Tris, 0.1 M NaCl, 36% w/v polyethylene glycol; $\circ$, 25 mM Tris, 0.05 M NaCl, 62% w/v polyethylene glycol; $\square$, 10 mM potassium phosphate buffer, 36% w/v polyethylene glycol, the medium used for the X-ray studies. The buffer concentrations in the crystal are higher than in the surrounding medium (by about a factor of 1.4 for 36% polyethylene glycol and a factor of 2 for 62% polyethylene glycol), assuming that polyethylene glycol is not contained in the crystal. The pH in the interior of the crystal is assumed to be the same as the measured pH of the buffered polyethylene glycol medium.



the partial relief of strain in the liganded structure. Haemoglobin in solution with strong allosteric effectors bound has a T-state affinity comparable with that of haemoglobin in the crystal^{27,28}, suggesting that it also has a liganded conformation similar to that of the crystal and will exhibit a much reduced or absent tertiary Bohr effect²⁶.

Some tertiary conformational changes on ligand binding do, however, occur in the crystal¹⁹. As for crystals of haemoglobin in the R quaternary structure^{9,29,30}, the decrease in polarization ratio on oxygenation (Fig. 1b) indicates a change in β haem orientation (the X-ray structure shows that there is very little change in the α haem orientation). The calculated β haem tilt (A. M. *et al.*, unpublished observation) is very close to that found for carbon monoxide binding in the T quaternary structure of the metal hybrid, $\alpha(\text{Ni})_2\beta(\text{FeCO})_2$ ³¹. This finding suggests that the tertiary conformational change in β subunits is very similar in the crystal for oxygen and carbon monoxide binding. These conformational changes do not, however, generate any cooperativity or Bohr effect.

The final point to discuss is the difference between the affinities of the α and β subunits. In the electron density map calculated from the X-ray data, the binding site for the α haems was fully occupied by an oxygen molecule, whereas no oxygen molecule was bound to the β haems¹⁷⁻¹⁹. This result predicts a distinctly biphasic binding curve (provided that binding is noncooperative). By contrast, the observed binding curve is monophasic. Furthermore, the linearity of the Hill plots (Fig. 2) indicates that the α and β binding constants differ by less than a factor of 3, compared with a factor of at least 15 predicted by the X-ray result (assuming an uncertainty of 20% in the occupancies of the oxygen binding sites³²). Also, the α and β haems have different orientations with respect to the a and c crystal axes, and therefore contribute differently to the absorption. The $p50$ calculated from the a -axis data, where the α and β haems contribute equally to the absorption, is about 5% higher than for the c -axis data, where the α haems contribute 55-60% of the absorption. We can thus estimate that the α haems have about double the affinity of the β haems, which is similar to what is observed in solution for the T state both in the presence and in the absence of inositol hexaphosphate^{33,34}. Could the discrepancy between the X-ray and optical studies in the fractional oxygenation of the β haems be due to the fact that the oxygen molecules are 'invisible' in the electron density map because of disordering of the distal oxygen atom? To answer this question will require the determination of the X-ray structure of crystals for which the fractional saturation with oxygen and fractional oxidation are known from their optical absorption spectra. □

Received 18 February; accepted 19 April 1991.

1. Monod, J., Wyman, J. & Changeux, J. *J. molec. Biol.* **12**, 88-118 (1965).
2. Perutz, M. F. *Nature* **228**, 726-739 (1970).
3. Antonini, A. & Brunori, M. *Hemoglobin and Myoglobin in their Reactions with Ligands* (North-Holland, Amsterdam, 1971).
4. Szabo, A. & Karplus, M. *J. molec. Biol.* **72**, 163-197 (1972).
5. Shulman, R. G., Hopfield, J. J. & Ogawa, S. *Q. Rev. Biophys.* **8**, 325-420 (1975).
6. Edelstein, S. J. *A. Rev. Biochem.* **44**, 209-232 (1975).
7. Baldwin, J. M. *Prog. Biophys. molec. Biol.* **29**, 227-320 (1975).
8. Ackers, G. & Smith, F. R. *A. Rev. Biophys. Biochem.* **16**, 583-609 (1987).
9. Perutz, M. F., Fermi, G., Luisi, B., Shaanan, B. & Liddington, R. C. *Accs Chem. Res.* **20**, 309-321 (1987).
10. Perutz, M. F. in *The Molecular Basis of Blood Diseases* (eds Stamatoyannopoulos, G., Nienhuis, A. W., Leder, P. & Majerus, P. W.) 127-178 (Saunders, Philadelphia, 1987).
11. Perutz, M. F. *Q. Rev. Biophys.* **22**, 139-236 (1989).
12. Rossi, G. L. & Bernhard, S. A. *J. molec. Biol.* **49**, 85-91 (1970).
13. Mozzarelli, A., Peracchi, A., Rossi, G. L., Ahmed, S. A. & Miles, E. W. *J. biol. Chem.* **264**, 15774-15780 (1989).
14. Haurowitz, F. *Z. physiol. Chem.* **254**, 266-274 (1938).
15. Perutz, M. F. *Acta Crystallogr.* **6**, 859-864 (1953).
16. Perutz, M. F., Bolton, W., Diamond, R., Muirhead, H. & Watson, H. C. *Nature* **203**, 687-690 (1964).
17. Brzozowski, A. *et al. Nature* **307**, 74-76 (1984).
18. Liddington, R. C. thesis, Univ. York (1986).
19. Liddington, R., Derewenda, Z., Dodson, G. & Harris, D. *Nature* **331**, 725-728 (1988).
20. Lee, A. W., Karplus, M., Poyart, C. & Bursaux, E. *Biochemistry* **27**, 1285-1301 (1988).
21. Mills, F. C., Johnson, M. L. & Ackers, G. K. *Biochemistry* **15**, 5351-5362 (1976).
22. Imai, K. *Allosteric Effects in Haemoglobin*. (Cambridge University Press, Cambridge, 1982).
23. Gill, S. J., Di Cera, E., Doyle, M. L., Bishop, G. A. & Robert, C. H. *Biochemistry* **26**, 3995-4002 (1987).

24. Imai, K. & Yonetani, T. *J. biol. Chem.* **250**, 2227-2231 (1975).
25. Chu, A. H., Turner, B. W. & Ackers, G. K. *Biochemistry* **23**, 604-617 (1984).
26. Kilmartin, J. V., Anderson, N. L. & Ogawa, S. *J. molec. Biol.* **123**, 71-87 (1978).
27. Lalezari, I. *et al. Biochemistry* **29**, 1515-1523 (1990).
28. Marden, M. C., Bohn, B., Kister, J. & Poyart, C. *Biophys. J.* **57**, 397-403 (1990).
29. Makinen, M. W. & Eaton, W. A. *Nature* **247**, 62-64 (1974).
30. Eaton, W. A. & Hofrichter, J. *Meth. Enzym.* **76**, 175-261 (1981).
31. Luisi, B., Liddington, R. C., Fermi, G. & Shibayama, N. *J. molec. Biol.* **214**, 7-14 (1990).
32. Brzozowski, A. *et al. in Hemoglobin* (eds Schnek, A. G. & Paul, J.) 37-51 (Editions de l'Université de Bruxelles, Belgium, 1984).
33. Shibayama, N., Morimoto, H. & Miyazaki, G. *J. molec. Biol.* **192**, 323-329 (1986).
34. Sawicki, C. A. & Gibson, Q. H., *J. biol. Chem.* **252**, 7138-7147 (1977).
35. Gill, S. J. *Meth. Enzym.* **76**, 427-438 (1981).

ACKNOWLEDGEMENTS. We thank A. Szabo for helpful discussions, E. Padlan for X-ray diffraction measurements, M. Brunori and M. Perutz for critical reading and comments on the manuscript, and B. Luisi and R. Liddington for providing unpublished atomic coordinates. This work was supported in part by a binational grant and by the Target Project on Biotechnology and Bioinstrumentation of the National Research Council of Italy.

CORRECTIONS

Negative regulation of human *c-fos* expression by the retinoblastoma gene product

Paul D. Robbins, Jonathan M. Horowitz
& Richard C. Mulligan

Nature **346**, 668-671 (1990)

THIS letter in the 16 August 1990 issue of *Nature* reported that the protein product of the retinoblastoma susceptibility gene (RB-1) can repress *c-fos* expression and AP-1 transcriptional activity in both serum-induced and cycling 3T3 cells. We also identified a *cis*-acting element in the human *c-fos* promoter that can confer repression by Rb to a heterologous promoter. We now report that the intact *c-fos* promoter and the deletion mutants of the *c-fos* promoter used in the study were incorrectly described as being of human origin. They are, in fact, of mouse origin. The oligonucleotide RCE element, however, is indeed of human origin as stated. Aside from the erroneous description of the promoter sequences, the data and conclusions of the study remain as reported. Our recent unpublished studies confirm that, as expected, an intact human *c-fos* promoter (which contains the RCE sequences identified in our report) is negatively regulated by Rb, and that an oligonucleotide representing mouse sequences analogous to the human RCE (-100 to -62) also confers negative regulation by Rb upon the TK promoter. We regret the confusion that the error has caused.

Action potentials must admit calcium to evoke transmitter release

Rosel M. Mulkey & Robert S. Zucker

Nature **350**, 153-155 (1991)

IN the above letter in the 14 March issue, the number given for the increase in frequency of miniature e.p.s.ps due to DM-nitrophen photolysis in normal Ringer's (1,700-22,500%) is incorrect and should read 1,700-12,500 Hz (middle of the left column, page 155).

Validatable virus removal from protein solutions

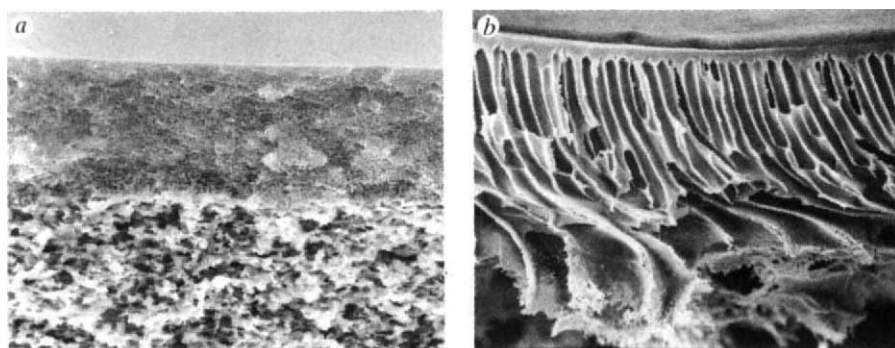
Anthony J. DiLeo and Anthony E. Allegrezza Jr

A membrane-based filtration system, designed for use in the production of biotherapeutics, predictably removes viruses as a function of size.

VIRAL contamination is a major concern in the manufacture of biopharmaceuticals. Continuous cell culture, which is used increasingly for producing recombinant proteins, offers an opportunity for viral multiplication and infection to occur. Most rodent cell lines have been found to be contaminated with retrovirus or retrovirus-like particles. The cell culture medium (often containing mammalian serum) and other process reagents can also introduce viruses, and although presumably rare, breaches in Good Manufacturing Practices are a potential source of viral contamination. Monoclonal antibodies (mAbs) are administered for a growing number of medical purposes and are also very useful in the purification of various therapeutic proteins. They can become contaminated with viruses introduced by the myeloma cells and natural antibodies that are used to create the mAb-producing hybridomas.

Viral contamination is problematic as current techniques for both the detection and the inactivation/removal of viruses leave much to be desired. Viral assays lack the sensitivity to detect titres which, although low, may still be of medical concern. A specific assay must be performed for each particular virus and there is no method of determining the total virus titre. To counter viral contamination, it is typical for commercial operations to use concurrently several approaches, each based on a different physical or chemical principle. The clearance factors of each approach are taken to be cumulative and theoretical clearance factors of 12 orders of magnitude or greater are routine. Unfortunately, chemical, thermal, ultraviolet, and gamma-ray inactivation treatments do not affect all viruses equally. And as these physical or chemical methods allow the inactivated virus to remain with the product, recombinant incorporation of viral DNA/RNA is a concern. With some techniques, protein stabilization is required and, of course, any viricidal and/or stabilizing chemicals introduced during the manufacturing process must be reduced to safe levels in the final product.

Physical separation methods have been limited to chromatography and ultrafiltration (UF); ultracentrifugation uses such high forces to separate viruses from proteins that it is impractical to scale up to process typical production volumes. Chromatographic methods (primarily ion exchange and gel



Scanning electron micrographs of the cross-sectional structure of a, Viresolve/70 membrane and b, a typical ultrafiltration membrane (PTMK, Millipore).

permeation) are often used but they exhibit a wide range of clearance factors depending on the nature of the virus.

Ultrafiltration is also used¹, but validation must be done on a case-by-case basis because of unproven membrane consistency and adsorption effects. A recent study³ indicates that typical polysulphone and nylon membranes do not retain particles solely by sieving, as has been generally assumed. The data show that viral retention does not increase with the size of the particles, indicating a significant percentage of openings in the membrane surface that pass all particles. Furthermore, these membranes show decreased viral retention when previously treated with albumin. Such results suggest that adsorption of virus particles on the surfaces of these membranes is their primary mechanism of viral retention. Overall, the use of presently available UF membranes for the removal of viruses is limited by the extent of defects or very large pores formed on the active skin of the membrane.

Uniform pore size

A new type of filtration membrane, Viresolve (Millipore, Bedford, Massachusetts), has been developed that is designed specifically for the removal of viruses from protein solutions. Its structure differs fundamentally from typical UF membranes (see figure): Viresolve consists of a thin (10 µm) porous membrane formed on a preformed microporous membrane, while UF membranes are generally 100–150-µm thick and supported on a woven or non-woven fabric. The combination of using a microporous membrane as the substrate and forming a thin membrane, particularly one without the macrovoids ('fingers') usually found in UF membranes, produces a new type of membrane with enhanced virus rejection properties. Adsorption plays little or no role in virus retention because the polyvinylidene fluoride material of the Viresolve membrane is rendered hydrophilic with hydroxy-propyl acrylate and so exhibits very low protein binding.

THEORETICAL MINIMUM VIRUS REMOVAL USING VIRESSOLVE/70 SYSTEM

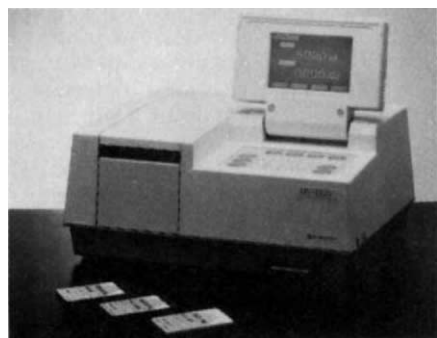
Virus	Diameter (nm)	One-stage (log removal)	Two-stage (log removal)
Parvovirus	22		2.7–3.0
Non-A, non-B (hepatitis virus)	27		3.7–4.0
Cocksakie virus A	30		4.3–4.6
Bovine diarrhoea virus	40	3.3–3.6	6.5–6.7
Hepatitis B virus	42	3.5–3.8	6.8–7.1
Human papilloma virus	55	4.8–5.1	9.4–9.7
Adenovirus	70	6.0–6.3	
Reovirus 3	75	6.3–6.6	
Marburg & Ebola	80	6.6–7.0	
Human immunodeficiency virus	100	7.4–7.7	
Herpesvirus	100	7.4–7.7	
Lymphocytic choriomeningitis virus	110	7.4–7.8	

Calculated minimum log removal factors for one- and two-stage systems processing 10-litre batches of virus in buffer only. Typical protein recovery with the parameter values used is > 94%.

run initiation to data storage. The monochromator provides light over a wavelength range of 190–800 nm. When used as an analytical ultracentrifuge, the system must be used with the An-60 Ti analytical rotor, which features three positions for samples and a fourth for calibrating radial distances. The monochromator can be installed and removed in minutes and so the XL-A can easily be converted to a preparative ultracentrifuge. The Optima XL-A costs \$103,000 (US), which includes the start-up package. Beckman will be in hall 6.2, stand BC8–10.

Overlay is the new **graphics program** from Spectra-Physics, which can be used for the direct qualitative comparison of up to 20 chromatograms (*Reader Service No. 105*). The program is designed to run in the Spectra-Physics WINer, WINer/286 and WINer/386 workstation environments. Using Overlay, chromatograms can be drawn directly on top of one another, or offset in the X and/or Y direction; height and/or time can be 'normalized' to compare entire chromatograms or sections of chromatograms; small differences in chromatograms can be identified by manipulating traces by addition or subtraction; and the signal-to-noise ratio can be enhanced to facilitate the identification of trace components. The Overlay program is available on 3.5- and 5.25- inch diskettes for use with any Spectra-Physics PC-based system. See Spectra-Physics in hall 6.3, stand FG14–15.

Stop by the Shimadzu display, hall 6.3, stand D20–21/E21, and see the company's latest **UV/visible spectrophotometer**, the



Shimadzu's UV-1201 spectrophotometer.

ADVERTISEMENT

Severn Biotech Ltd.

Custom DNA Synthesis
Probes • Primers • Oligo's

*For a fast efficient
Cost effective service*

BUY SEVERN BIOTECH

Tel: 0562 825286 Fax: 0562 825284
Unit 23, Hoobrook Enterprise Centre,
Worcester Road, Kidderminster DY10 1HY

UV-1201 (*Reader Service No. 106*). The UV-1201 provides an ultrahigh scanning speed of 6,000 nm min⁻¹. In addition, the optical system has a high energy throughput, with 0.05 per cent stray light at 340 nm; this allows measurements in the wavelength range of 200–1,100 nm, says Shimadzu. Time-course and kinetic data, and experimental parameters can be stored on 'smart' IC-cards. Application programs for spectrum presentation, quantitation, time-course, or kinetics are available on program-pack cards, in German, English, French, Italian, Spanish and Japanese.

All about image

Image processing and computer hardware, a selection of software modules and a relational data base make up Ciba Corning's Tracktel **image processing system** for elec-



Tracktel imaging system — see Ciba.

trophoresis, densitometry and nucleic acid sequencing (*Reader Service No. 107*). After capturing the image on the electrophoretic plate, Tracktel will calculate molecular weights, band intensities and compare individual tracks with those already stored in the relational data base. The stored images form a permanent record of a gel, which is non-degradable and which can be archived. A selection of software modules for forensic identification, paternity testing or laboratory applications allow users to tailor the system to suit individual needs. Tracktel can accommodate autoradiograms, membranes, photographs and gels (including fluorescent gels). Catch Ciba Corning in Hall 6.2, stand J12.

UVP International introduces the GDS2000 **gel documentation system**, which has been specifically designed for the photography of low-light fluorescent patterns as in the case of DNA gels stained with ethidium bromide (*Reader Service No. 108*). In addition, the system can be used for the photography of other materials such as protein gels and TLC plates. Mini gels and those up to 20 × 20 cm can be accommodated. The GDS2000 system consists of a camera with IR and UV filters, power supply unit for the camera, monitor, video printer, camera and system stand. Images, either full-gel images or one-dimensional 'track images' can be exported as PostScript files for printing on a laser printer or incorporation into a desktop publishing package. The GDS2000 system can be further enhanced with the addition of UVP's SW2000 soft-

ware analysis package. This package allows the user to perform a variety of functions, including image enhancement, band profiling, molecular weight and base-pair calculations. The software is compatible with IBM PCs or compatibles. With different frame-store hardware, the package could also run on VAX, Sun or Silicon Graphics workstations. UVP will be in hall 6.3, stand E12A.

Histochemical helpers

Neslab — see hall 6.3, stand B15 — has introduced a handy-sized **tissue freezing bath** for the quick freezing of tissues and other biological samples (*Reader Service No. 109*). Immersion in the Histo Bath freezes tissues and prevents the formation of ice crystals, which sometimes form in samples frozen using a cryostat, says Neslab. Rapid cooling to -60 °C and stainless steel wetted parts are features of the Histo Bath. The unit is supplied with a stainless steel sample basket and a gasketed PVC cover, which is designed to eliminate bath fluid ice-up.

The Varistain 24-4 **automatic slide stainer** from Shandon Scientific can be programmed to perform up to six separate histology or cytology staining routines (*Reader Service No. 110*). The slide stainer is based on a carousel design and features a turntable that holds the 24 reagent troughs, each with a 750-ml capacity. Users are provided with an option of three 24-step programmes or six 12-step staining programmes. The stainer can accommodate up to 64 vertically-positioned microscope slides in each of two slide carriers. Programming is achieved by selecting an immersion time for a series of numbered

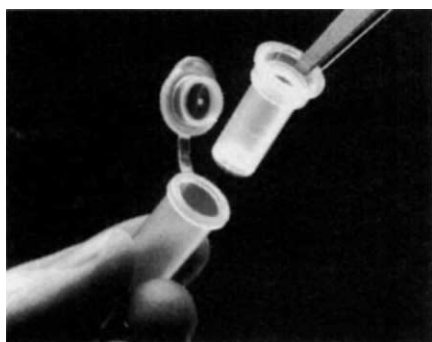


Varistain carousel slide stainer.

reagent stations (or troughs), which can be anything from one second to 59 minutes 59 seconds for each station. Users can stop the stainer and check on the slides at any time during the run without affecting the automatic cycle, says Shandon. The Varistain stainer is supplied with staining and wash troughs, outlet tubing and fittings, and two 64-position slide containers. See hall 6.1, stand G4.

Molecular minutes

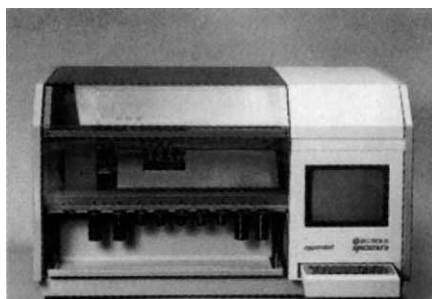
Millipore has developed a new Ultrafree-Probind **centrifugal filter unit** for nucleic acid purification, which eliminates the need



Centrifugal filter unit.

to perform phenol/chloroform extractions (*Reader Service No. 111*). The polyvinylidene fluoride (PVDF) membrane housed within the Ultrafree-Probind unit exhibits preferential binding of protein over double-stranded DNA, says Millipore. The filter unit consists of a 400- μ l sample cup with the PVDF membrane sealed to the bottom. The entire unit fits inside a 1.5- μ l microcentrifuge tube and so is designed for use in any microcentrifuge at spin speeds of 12,000 g. See hall 5.1, stand DE8-11.

Synostat D is the new automated DNA synthesizer from the Molecular and Cell Biology Division of Eppendorf, which is capable of synthesizing up to three oligonucleotides simultaneously (*Reader Service No. 112*). The synthesis scale of the instrument ranges from 0.05 to 15 μ mol. On-line

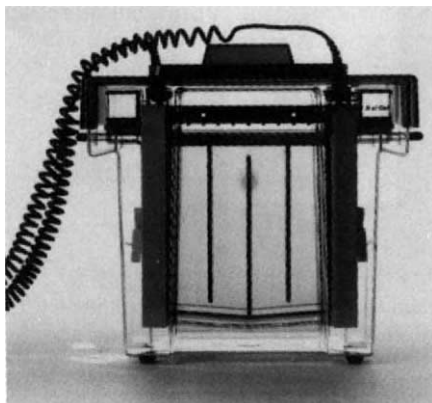


DNA synthesis from 0.05-15 mols.

optosensors monitor reagent flow and filling of the synthesis columns. Phosphoramidite reagents in powdered form are dissolved directly on the instrument. An accurate metering syringe in combination with a patented zero dead-volume valve block, which can be serviced by the user, ensure the economical use of expensive nucleotides, says Eppendorf. In addition, programmable syringe delivery of nucleotides allows controlled mixed-base positions. For high-throughput, the synthesizer can be programmed for cycle times of less than four minutes. See hall 5 ZG, stand F21-23.

A cooling core with a raised electrode for safer operation, silicone-coated power cords with flexible jack sleeves and a design that ensures leak-free gel casting without grease or agarose plugs are features of the new Protean II xi electrophoresis cell from Bio-Rad (*Reader Service No. 113*). The cell can hold both 16- and 20-cm slab gels, or up to 16

tube gels; the extra run distance that is provided with a 20-cm gel can often be useful in resolving bands or 2-D spots of similar molecular weights, says Bio-Rad. Temperature regulation not only provides added flexibility with experimental times but allows the cell to maintain a temperature of 4 °C for native or preparative applications or 50 °C for special DNA runs. With cooling, Bio-Rad claims that SDS-PAGE runs can be

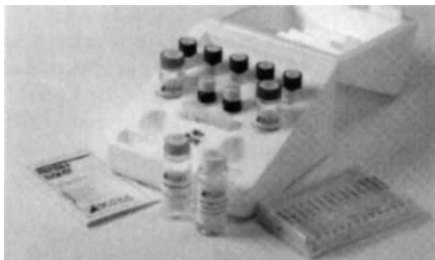


Temperature-controlled electrophoresis cell.

completed in less than three hours. Bio-Rad will be in hall 6.3, stand G5-7.

Ideas in immunodetection

Biotage has introduced a Protein A ELISA kit that can be used for the quantitative analysis of Protein A in wash fluids and immunoglobulin samples purified by solid-phase Protein A methods (*Reader Service No. 114*). The assay, which is based on a non-isotopic procedure, can be completed in less than four hours at room temperature and is sensitive to 2 ng ml⁻¹, says Biotage. The kit contains one 96-well strip plate coated with anti-Protein A antiserum, as well as all the



Protein A ELISA in less than 4 hours.

necessary standards, reagents, diluents and buffers in premeasured amounts — sufficient reagents to perform four independent assays including standard curves. Biotage, a manufacturer of low-pressure and high-pressure LC systems, will be in hall 5 ZG, stand B2.

Ready-to-use chloro-naphthol-containing peroxidase substrate for blotting applications, which combines both substrate and chromogen in one solution, has been introduced by Janssen (*Reader Service No. 115*). Poor solubility and poor stability of most chromogens and peroxidase has, in the past, made it necessary to prepare substrate solutions immediately before use. PrestoSol BL,

ADVERTISEMENTS

METABOLITES^{analysed} rapidly, enzymatically, using micro-samples

with
the



ANALOX GM7 MICRO-STAT

A wide range of biological fluids including whole blood are acceptable for direct injection. Turbidity no problem. Sample size 2.5 - 25 μ l. Results in 20 seconds.

glucose	lactate	urea
alcohol	pyruvate	creatinine
cholesterol	ammonia	3-OH butyrate
uric acid		acetoacetate

Analox Instruments Ltd

8 Goldhawk Industrial Estate
Brackbury Road Tel: 081-749 7635
Hammersmith, London W6 0BQ Fax: 081-740 6608



The original SINGER MK I MICROMANIPULATOR single or in pairs

- FOR
- XENOPUS OOCYTE INJECTION
 - PIPETTING OF CELLS IN SUSPENSION
 - MICRODISSECTION AND MICRO SURGERY

FAX: +44.984.41166 for details
(24 hours)

Custom DNA

Purified and Shipped in 48 hours,
Worldwide.

100 μ g, \$10.00 a base for the first 15 bases,
\$7.50 for each additional base.
Delivered.

Research Genetics

1-800-533-4363

In the United Kingdom 0-800 89-1393
In Canada 1-800-533-4363

FAX 205-536-9016

Synthetic Oligonucleotides

Custom made probes and primers,
synthesised, purified and shipped in 3 days

£3.50 per base

(All inclusive)

0.2 μ M scale (approx. 500 μ g)

Modified bases, biotin or digoxigenin labelling,
data base analysis and full design facilities

Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX

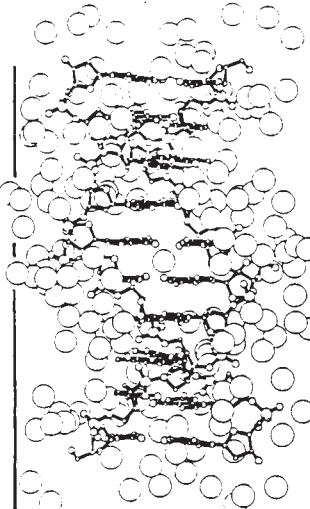
Tel. XX-44-(0)71 - 326 3126

FAX XX-44-(0)71 - 733 3877

ADVERTISEMENT

DNA CUSTOM SYNTHESIS

TIB MOLBIOL
BERLIN GmbH i.G.



OLIGONUCLEOTIDE
SYNTHESIS SERVICE

0.2 µM scale 5,-/base
 + 95,-/oligo

MODIFIED BASES

dl, dU, 5-J-dU, 5-Br-dU, 5-J-dC
5-Br-dC, 5-Met-dC, N6-Met-dA
 30,-/base

ANTISENSE

5'-Psoralen 200,-/oligo
Phosphothioate 30,-/base

LINKER

5'-Aminolinker, Thiolinker, Biotin
5'-Fluorescein, Phosphorylation
 200,-/oligo

PURIFICATION
BY GEL FILTRATION

SHIPPED WITHIN 2-3 DAYS
EXPRESS DELIVERY

All prices are in DM

Custom Peptide Synthesis
Custom Gene Synthesis
Custom Mutagenesis

ORDER PER FAX
+49 30 6261960

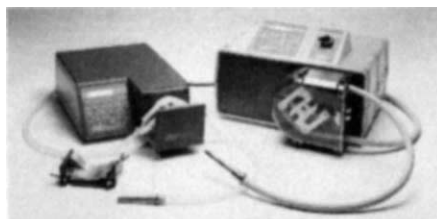
TIB MOLBIOL
EMSER STR. 103
D-1000 BERLIN 44

Reader Service No.86

however, has a shelf-life of at least 12 months and contains a high amount of chromogen concentrated in solution, says Janssen. See hall 6.1, stand D14.

Liquid assets

Watson-Marlow will be showing off its high-flow **digital pump**, the 605Di/R, for the first time at Achema — see hall 6.1, stand C12-13 (*Reader Service No. 116*). The pump, which is designed to dispense one litre in less



Pump it up — see hall 6.1, stand C12-13.

than four seconds, can be programmed with dosing and dispensing routines or, alternatively, can be operated remotely by analog or digital signals. Pump commands are entered via a membrane keypad and 32-character menu. The user enters a choice of calibration, dosing or auto control programmes. The calibration programme establishes the actual flow rate for a given tube size and is recommended for accurate dispensing. When using the dosing programme, the user enters the tubing size to be used, the speed of the pump, the size of the dose required either by weight or volume, and the number of doses with the time interval between each dose. When under auto control, the pump can be controlled by a computer supplying analog signals or an RS232 digital signal. The 605Di/R is connected to the signal source via an 18-pin socket on the back panel.

A pre-treatment and post-treatment package, and two units for the **production of purified water** are included in the Elgastat Option range (*Reader Service No. 117*).



Pure water on tap by Elga.

Option 1, a pre-treatment package that comes in a wall- or bench-mounted unit, is designed to improve the performance of the distillation process by reducing the scaling-up process and improving the quality of the distillate. The system is based on reverse osmosis and removes up to 98 per cent of dissolved solids and over 99 per cent of bacteria, pyrogens and particulates from still feedwater, says Elga. On the other hand, Option 2 is a post-treatment unit that Elga has developed to maintain and improve the

quality of pre-purified, stored water. While continuous recirculation of the purified water maintains water quality by inhibiting bacterial growth, a composite cartridge removes both organic and inorganic impurities and a UV-photochemical reactor cell provides continuous disinfection. Options 3 and 4 provide a means of producing general-grade laboratory water from a potable water supply. A composite pre-treatment cartridge combines with reverse osmosis, ionic/organic removal and UV disinfection to produce up to 8 litres per hour of purified water, with an energy consumption that is less than one hundredth that of a comparable still, says Elga. Option 4 includes all of the features of the Option 3, but also has an internal pump for continuous recirculation. See hall 6.2, stand G13.

Labware line-up

Heraeus has introduced a new **vacuum oven** which is intended for drying porous substances with low thermal conductivity, for drying or storing hygroscopic products, or for tempering and outgassing plastics or paraffin-embedded components (*Reader Service No. 118*). The vacuum oven is available in two



Drying by vacuum processing.

sizes. Safety features include an overtemperature safeguard and two panes of safety glass — the inner pane, which is made of 15-mm thick tempered safety glass, to hermetically seal the vacuum chamber, and a second pane, also safety-screened, to provide additional protection against heat and implosion. See hall 4.2, stand FG17-19.

The new digital, mercury-free **electronic temperature sensor** from Sartorius is designed to provide precise temperature control of solutions during heating operations (*Reader Service No. 119*). The IKA ETS-D sensor gives a continual and alternate display of desired and actual temperatures. Once the desired temperature value has been set, the unit prevents further inadvertent adjustment in order to safeguard against overheating of both the medium and the hot plate. See hall 6.3, stand HJ19-21. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.